

THESIS / THÈSE

MASTER EN SCIENCES INFORMATIQUES

Aide au diagnostic médical synthèse des méthodes et applications

Deroy, Patrick

Award date:
1979

Awarding institution:
Université de Namur

[Link to publication](#)

General rights

Copyright and moral rights for the publications made accessible in the public portal are retained by the authors and/or other copyright owners and it is a condition of accessing publications that users recognise and abide by the legal requirements associated with these rights.

- Users may download and print one copy of any publication from the public portal for the purpose of private study or research.
- You may not further distribute the material or use it for any profit-making activity or commercial gain
- You may freely distribute the URL identifying the publication in the public portal ?

Take down policy

If you believe that this document breaches copyright please contact us providing details, and we will remove access to the work immediately and investigate your claim.

FACULTÉS
UNIVERSITAIRES
N.D. DE LA PAIX
NAMUR



INSTITUT D'INFORMATIQUE

AIDE AU DIAGNOSTIC MEDICAL
SYNTHESE DES METHODES ET APPLICATIONS

PATRICK DEROY

Promoteur : J. FICHEFET

Mémoire présenté en vue
de l'obtention du grade
de Licencié et Maître
en Informatique.

ANNEE ACADEMIQUE 1978-1979

FACULTES
UNIVERSITAIRES
N.-D. DE LA PAIX
NAMUR

Bibliothèque

FM B 16

1979 / 16

FM B 16 / 1979 / 16



AIDE AU DIAGNOSTIC MEDICAL
SYNTHESE DES METHODES ET APPLICATIONS

PATRICK DERROY

Promoteur : J. FICHEFET

Mémoire présenté en vue
de l'obtention du grade
de Licencié et Maître
en Informatique.

US 3212933



6520 - 27021

A mon père.

Je remercie Monsieur le Professeur J. FICHEFET
pour son assistance et ses conseils nécessaires à l'élaboration de
mon mémoire.

Je tiens également à remercier le Professeur M. ROUX dont
les recherches sur le dossier médical informatisé et plus particulièrement
sur l'aide au diagnostic m'ont permis de réaliser de travail.

Je remercie aussi toute l'équipe du Laboratoire de
Biomathématiques, Statistiques et Informatique médicale du C.H.U. Timone
de Marseille, de m'avoir intégré en son sein durant mon stage et de
m'avoir prodigué les nombreux conseils utiles pour la rédaction de
mon mémoire.

TABLES DES MATIERES

	pages
<u>INTRODUCTION</u>	1
A. Introduction générale.	2
B. Présentation des chapitres ultérieurs.	7
 <u>1ERE PARTIE : SYNTHESE DES METHODES D'AIDE AU DIAGNOSTIC EXISTANTES</u>	 10
Chapitre I : Méthodologie.	11
1. Le diagnostic global.	12
1.1. Les méthodes d'analyse des données multidimensionnelles.	12
1.2. Les méthodes de discrimination non paramétriques.	17
2. Diagnostic pas à pas.	18
2.1. Les méthodes logiques.	18
2.1.1. Application du calcul booléen.	18
2.1.2. Extension à l'algèbre ternaire.	29
2.1.3. Démonstration automatique.	30
2.2. La méthode probabiliste.	31
2.3. Théorie des sous-ensembles flous.	36
Chapitre II : Application et études des résultats de quelques méthodes vues au chapitre I.	 38
1. Application à la pathologie thyroïdienne de la méthode basée sur la logique booléenne.	 40
2. Application à la pathologie thyroïdienne de la méthode basée sur la logique floue.	 46
2.1. Introduction aux ensembles Φ -flous.	46
2.2. Mise en oeuvre des fonctions Φ -floues.	50
2.3. Quelques notions sur l'interprétation du résultat des fonctions Φ -floues.	 53
2.4. Résultats obtenus lors du passage en machine de ces fonctions Φ -floues.	 54
3. Application à la pathologie thyroïdienne de la méthode de type multidimensionnelle.	 56
4. Application à la pathologie thyroïdienne de la méthode par pondération des caractères qualitatifs.	 63
5. Conclusion sur ces différentes applications.	67

<u>IIEME PARTIE : APPLICATION BASEE SUR LE TRAVAIL DE STAGE</u>	68
A. Etude du dossier médical : SYCVAR	69
Chapitre I : Introduction.	71
Chapitre II : Présentation du système.	74
2.1. Informations retenues.	75
2.2. Structure logique - dossier de base	76
- dossiers satellites.	
2.3. Entrée des informations.	77
2.4. Exploitations.	78
Chapitre III : Spécifications plus techniques à propos du système SYCVAR.	80
3.1. Généralités.	81
3.1.1. Les informations élémentaires.	81
3.1.2. Notions d'existence de l'information élémentaire.	81
3.1.3. Hiérarchie des informations.	82
3.1.4. Fichiers descripteurs pointés par l'intermédiaire du répertoire.	82
3.2. Notions d'implémentation de ces fichiers.	84
Chapitre IV : Utilisation du système.	89
4.1. Saisie des données.	90
4.2. Exploitation des données.	93
4.3. Facilité d'utilisation.	95
B. Aide au diagnostic par la méthode de démonstration automatique de théorèmes	96
Chapitre I : Notions élémentaires de logique	97
1.1. Logique propositionnelle	98
1.2. Logique du premier ordre.	99
1.3. Théorème de Herbrand.	101
1.4. Méthode de Davis et Putnam.	105

Chapitre II : Application de la logique au problème médical.	107
2.1. Utilisation de la méthode de Davis et Putnam sur un exemple simple de diagnostic.	108
2.2. Utilisation de la méthode de Davis et Putnam sur la pathologie thyroïdienne.	111
2.3. Etude des programmes utilisés.	115
2.3.1. ENTRAX : Programme d'acquisition des systèmes d'axiomes	115
2.3.1.1. Axiomes à définir.	115
2.3.1.2. Organigrammes et spécifications.	116
2.3.2. DEMONS : Programme d'acquisition des signes du patient et de diagnostic.	127
2.3.2.1. Prise en considération des signes du patient.	127
2.3.2.2. Démonstration par la méthode de Davis et Putnam.	129
2.3.2.3. Amélioration possible du programme de démonstration.	140
2.3.2.4. Applications du programme de démonstration.	143
<u>CONCLUSION</u>	151
<u>ANNEXES</u>	154
Annexe I Exemple de questionnaire relatif à la pathologie thyroïdienne.	155
Annexe II Exemple de résumé avec diagnostic obtenu par la méthode booléenne.	161
Annexe III Exemples de diagnostics obtenus par le programme de démonstration automatique de théorèmes.	164
<u>BIBLIOGRAPHIE</u>	166

°
° °

INTRODUCTION

A. INTRODUCTION GENERALE

Le domaine de l'informatique médicale était, voici 15 ans encore, relativement vierge et une véritable recherche s'imposait dans ce domaine. Comme toute recherche n'est rentable qu'à terme, ce n'est qu'aujourd'hui que l'on peut en connaître les premiers résultats.

Lorsque, il y a une dizaine d'années, l'on parlait d'informatique médicale, cela se limitait à une gestion administrative des informations dans les hôpitaux. En effet, l'hôpital est une grande entreprise à la gestion difficile. Le traitement des informations par ordinateur a dû y être introduit en raison de la croissance du nombre de lits, de la diversité des services; de l'accroissement et la diversification du personnel. C'est à ce niveau qu'est apparu l'informatique dite hospitalière : gestion des lits, gestion des salles spécialisées, gestion du personnel et surtout gestion comptable, administrative et financière.

C'est petit à petit que l'on est arrivé à créer des dossiers personnels pour les patients. Ces dossiers n'avaient qu'une seule utilité au départ : être intégré dans la gestion administrative.

Il n'y avait qu'un pas à franchir pour en arriver à un dossier médical plus élaboré. Actuellement, ce pas est franchi dans certains hôpitaux. On est ainsi arrivé à un dossier médical complet qui rassemble toutes les informations recueillies sur un malade à l'occasion de chaque consultation ou hospitalisation : outre la situation administrative, les données fournies par un questionnaire (interrogatoire) et par les examens cliniques et/ou paracliniques.

L'utilité de ce dossier a également évolué avec le temps. Le dossier médical remplissait une fonction de dossier, c'est-à-dire d'un endroit où se trouvent toutes les informations relatives à un malade. Il s'agissait donc d'un aide-mémoire par patient que l'on consultait comme on consulte un dictionnaire pour un mot nébuleux.

Les choses ont évolué. Le dossier garde néanmoins cette fonction de base mais d'autres applications ont été élaborées grâce à l'informatisation de ces dossiers

- au niveau d'un patient : édition de résumé
aide au diagnostic
- au niveau de dossier : études statistiques
aide à la recherche médicale
enseignement de la médecine.

D'autres thèmes d'activité médicale et de santé ont bénéficié des techniques informatiques :

- La surveillance automatisée des grands malades (monitoring) et son contrôle par ordinateur. En effet, la surveillance de tels malades doit être à la fois permanente (on doit à tout moment pouvoir connaître l'état instantané des paramètres recueillis), rétrospective (il faut, en cas d'accident, pouvoir disposer immédiatement de l'ensemble des informations recueillies dans les minutes ou les heures précédentes) et aussi prospective (il faut, si possible, essayer de prévoir les tendances évolutives de tels ou tels symptômes).
- Le contrôle par ordinateur d'expériences de laboratoire.
- La classification de la connaissance et de la documentation médicale ainsi que sa constante mise à jour.

Malgré sa grande influence sur la vie de notre société et sur la vie médicale en particulier, l'informatique reste avant tout une TECHNIQUE.

Il apparaît que la force de pénétration de la technique informatique dans les activités de santé dépendait beaucoup moins de la nature propre des activités choisies pour être soumises à l'informatisation que de la capacité des médecins à les appréhender pour s'y engager eux-mêmes et y appliquer, en liaison étroite avec les informaticiens, les ressources de la technique informatique.

L'augmentation de la capacité de participation des médecins à l'informatisation peut être considérée comme la pierre de touche des progrès de cette informatisation.

Quelques facteurs importants semblent avoir guidé les choix de l'informatisation en médecine. Nous en retiendrons trois :

1. La valeur ajoutée qu'apporte l'informatisation.
2. L'intérêt économique de l'informatisation.
3. La capacité de participation des médecins et des personnels de santé à l'informatisation.

1. La valeur ajoutée qu'apporte l'informatisation.

Les caractéristiques propres aux ordinateurs telles que : grande capacité de mémorisation, exploitation simultanée de nombreux modèles et procédures logiques rapides, temps de réponse raccourci, permettent d'apporter une plus-value en médecine.

Mais tout en médecine n'est pas INFORMATISABLE, ou tout au moins ne l'est pas immédiatement. L'informatisation demande au médecin une remise en question de ses connaissances, une vision claire et précise des maladies à traiter par ordinateur. De plus un travail préliminaire de coopération médecin-informaticien est nécessaire afin de donner à chaque mot une signification propre.

Il n'est pas toujours nécessaire d'avoir des modèles logiques complexes pour obtenir des résultats satisfaisants. Nous verrons en effet, des modèles simples ne nécessitant pas l'emploi de gros moyens.

Le travail en temps réel n'est pas toujours nécessaire. Un travail avec léger différé permet souvent de grosses économies.

2. L'intérêt économique de l'informatisation.

La valeur ajoutée par l'informatique ne peut s'exprimer que si l'informatisation est économiquement rentable.

Cet intérêt économique doit être calculé par un groupe d'économistes et de médecins. Cet intérêt ne se traduit presque jamais par une réduction du coût global de l'activité informatisée. Il tient plutôt à une augmentation de l'efficacité qui réduit ainsi le rapport coût-efficacité.

3. La capacité de participation informatique des médecins.

La capacité de participation des médecins à l'informatisation est probablement le plus déterminant des trois facteurs appelés à guider les choix en informatique médicale. Cette capacité est très inégalement répartie, les facteurs de sollicitation paraissant avoir une moindre force que les facteurs de dissuasion.

a) Les facteurs de sollicitation.

L'introduction de l'informatique en médecine libérerait le médecin praticien de tâches subalternes ou triviales. La gestion informatique des contraintes administratives auxquelles il est soumis en allègerait le poids. L'aide informatique à la décision conforterait sa démarche et lui permettrait de mieux répondre aux interrogations qu'il se pose dans les domaines incertains de son savoir. Elle diminuerait son obligatoire allégeance aux médecins spécialistes. Ceux-ci n'auraient plus à intervenir que dans les cas particulièrement complexes.

Le temps ainsi gagné par le médecin dans sa pratique quotidienne favoriserait la reprise du dialogue avec le malade. Le médecin pourrait aussi consacrer ce temps à perfectionner son savoir.

b) Les facteurs de dissuasion.

La résistance des médecins à participer à l'informatisation de l'activité médicale constitue le principal obstacle au développement de cette informatisation. Cette résistance ne tient pas à leur manque de culture informatique. Ils parviendraient aisément, au prix d'un effort minime, à acquérir le minimum des connaissances nécessaires à l'usage de cette technique. Cette résistance tient avant tout à la crainte que l'informatique ne bouleverse l'ordonnement des activités médicales et la place de l'institution médicale dans la société. De plus, peu de médecins croient réellement à l'aide de la technique informatique dans des domaines aussi subjectifs que le diagnostic.

L'aliénation du secret médical par l'introduction de l'informatique en médecine est un autre motif fréquemment invoqué par les médecins pour justifier leur réticence à participer à l'aventure informatique en médecine. Il est vrai que la mise en mémoire d'ordinateur des dossiers médicaux de chaque

citoyen expose au risque qu'ils soient consultés par d'autres que des médecins. Au même titre qu'on est parvenu à préserver le secret bancaire, on ne voit pas pourquoi l'utilisation de clés spéciales ne parviendrait pas à préserver le secret médical. Les solutions techniques existent. Elles risquent toutefois d'être contournées par le "pouvoir" politique et administratif.

Un autre problème lié plus spécifiquement au diagnostic automatique est le partage des responsabilités d'un mauvais diagnostic sur les programmes élaborés par l'informaticien ou vice versa. Ceci risque donc de créer un malaise médical.

Nous avons ainsi fait un tour d'horizon de toutes les possibilités et problèmes relatifs à l'introduction de l'informatique en médecine. Nous avons également relaté les différents pôles d'application de l'informatisation médicale.

Il est impossible d'approfondir ici tous ces points, si bien que nous nous limiterons à un point particulier :

"L'AIDE A LA DECISION DIAGNOSTIQUE PAR ORDINATEUR"

Nous nous bornerons aux problèmes techniques du diagnostic automatique. Nous passerons sur les polémiques politiques, sociales ... engendrées par ce problème.

B. PRESENTATION DU TRAVAIL

Comme nous l'avons soulevé, le domaine couvert par l'informatique médicale est très vaste. C'est pourquoi nous avons dû restreindre notre champ de travail. Le sujet retenu a été :

"LE DIAGNOSTIC MEDICAL ASSISTE PAR ORDINATEUR"

Le travail présenté ci-après se découpera en deux parties principales :

- synthèse des méthodes existantes
- application basée sur le travail de stage - dossier SYCVAR
 - aide au diagnostic par la méthode de démonstration automatique de théorèmes.

1ERE PARTIE : SYNTHESE DES METHODES EXISTANTES

Le diagnostic médical repose, schématiquement, sur la confrontation des symptômes et signes observés sur un malade avec d'autres symptômes et signes connus dont les associations entraînent des décisions diagnostiques.

Chapitre I : Méthodologie

Dans cette synthèse d'assistance au diagnostic, nous avons pu regrouper ces méthodes en deux catégories.

- méthodes globales de reconnaissance.

Ces méthodes ont hérité de nombreux travaux des mathématiciens et statisticiens et il est donc normal qu'on se soit adressé à elles, dès les premières tentatives d'aide à la décision.

Dans "l'informatique médicale" publié dans la collection "Que sais-je", on montre assez bien à quoi correspond le diagnostic global.

"Dans une foule, nous reconnaissons un individu parmi beaucoup d'autres sans avoir à analyser séparément la forme de son visage, la couleur de ses yeux, le type de sa coiffure, etc... Tous ces éléments sont combinés par notre cerveau quasi instantanément."

Il s'agit donc d'une comparaison globale de l'ensemble des caractéristiques d'un individu avec l'ensemble des caractéristiques d'une maladie déjà connue qui amène au diagnostic.

Nous survolerons quelques-unes de ces méthodes dans cette partie du travail.

- méthodes "pas à pas"

Ces méthodes suivent plus exactement le processus de réflexion de l'esprit humain. Ce dernier ne peut pas prendre en compte simultanément plus de quelques signes. Ces signes qui sont constitués par des réponses à un questionnaire (ou des examens) déterminent le choix des questions suivantes (des examens complémentaires). On procède donc étapes par étapes, utilisant à chaque pas un nombre restreint de renseignements.

Le développement de la logique mathématique fournit un outil particulièrement bien adapté à l'aide à la décision en général et à la décision médicale en particulier.

Le calcul des probabilités basé sur le célèbre théorème de Bayes trouve également ici de nombreux débouchés.

Chapitre II : Applications et études des résultats

Les applications traitées ici ont été testées au CHU Timone de Marseille et ont fait l'objet d'études de performance. La pathologie traitée par ces applications est la pathologie thyroïdienne pour laquelle un dossier médical (SYCVAR) a été élaboré. SYCVAR représentera la connaissance médicale de base nécessaire à l'application de ces méthodes. (SYCVAR sera développé en 2ème partie)

Nous appliquerons une méthode de logique booléenne, de logique floue, une méthode discriminante, une méthode par pondération de facteurs qualitatifs et nous essayerons d'en tirer quelques conclusions.

2EME PARTIE : APPLICATION BASEE SUR LE TRAVAIL DE STAGE

Cette deuxième partie est spécifique au stage effectué au C.H.U. Timone de Marseille. Il s'agit d'une application basée sur la Pathologie thyroïdienne.

Chapitre I : Présentation du dossier médical : SYCVAR

Dans ce premier chapitre, il s'agit d'une étude du dossier médical : SYCVAR : système actuellement en exploitation à l'hôpital de la Timone.

Chapitre II : Aide au diagnostic par la méthode de démonstration automatique de théorèmes

Dans ce deuxième chapitre, nous avons voulu présenter une nouvelle méthode d'aide au diagnostic : décision au diagnostic par la méthode de :

DEMONSTRATION AUTOMATIQUE DE THEOREMES.

Les programmes relatifs à cette méthode ont été passés en machine (cii Mitra 125). C'est l'explication de la méthode, des programmes et des résultats obtenus qui constituera ce deuxième chapitre.

o

o . o

PREMIERE PARTIE

CHAPITRE I

MÉTHODOLOGIE

Les méthodes d'aide au diagnostic.

Avertissement.

Dans cette partie essentiellement théorique de notre étude, nous passerons en revue les principes inhérents aux méthodes actuellement appliquées dans certaines applications d'aide au diagnostic. Ce recueil ne se veut pas exhaustif mais il tente de présenter de façon plus ou moins objective ce qui a été découvert au cours des stages. Aussi avons-nous décidé d'uniformiser la présentation de ces méthodes. Ce chapitre est donc le fruit d'un travail commun à Th. Biernaux (*) et P. Deroy.

§ 1. Le diagnostic global.

1.1. Les méthodes d'analyse multidimensionnelle.

Ces techniques présentent un certain intérêt pour :

- la mise en évidence de certains aspects de l'association entre variables,
- l'étude des relations entre un paramètre (ou plusieurs) et un ensemble d'autres paramètres.
- la partition d'un ensemble d'observations en groupes homogènes,
- la détermination du "si" et du "comment" deux ou plusieurs groupes différents dans leur profil multidimensionnel.

La notion fondamentale de toutes ces techniques est celle de distance servant à définir la proximité ou la ressemblance existant entre observations ou paramètres.

Cette notion de distance distingue différentes méthodes d'analyse multidimensionnelle, parmi lesquelles il nous faut citer :

- l'Analyse des Correspondances,
- l'Analyse en Composantes Principales,
- l'Analyse Discriminante Linéaire ou Quadratique.

(*) T. Biernaux : "AIDE AU DIAGNOSTIC MEDICAL - Conception et analyse d'un système d'information à l'usage du Centre Anti-poisons de Bruxelles".

Comme le développement théorique de ces principales méthodes applicables ne présente que peu d'intérêt dans le cadre de notre étude, nous renvoyons le lecteur plus intéressé à trois ouvrages :

- BENZECRI J.P. 'Analyse des Correspondances'
tome II DUNOD 1973.
- DRAPER N.R. SMITH H. 'Applied Regression Analysis'
WILEY 1967.
- SEBESTYEN G.S. 'Decision Making Process in Pattern Recognition'
THE MAC MILLAN COMPANY 1962.

principe :

Ces méthodes procèdent généralement en deux étapes. La première utilise un échantillon de malades de diagnostic connu dont elle analyse les observations et parfois même les paramètres (Analyse des Correspondances). La seconde étape prend en compte un nouveau cas clinique qu'elle confronte à la connaissance scientifique tirée de l'échantillon et synthétisée, en vue d'en établir le diagnostic ou le pronostic.

1.1. a. Méthodes descriptives.

La méthodologie de l'analyse factorielle est basée sur la recherche d'un système réduit de références généré par de nouveaux axes. Ces nouveaux axes, appelés axes factoriels, sont obtenus par une rotation dans l'espace à p dimensions du système de coordonnées initial dans une orientation correspondant aux directions de variance maximum. Les axes factoriels sont choisis en ordre décroissant.

Il suffit alors de ne considérer que les axes factoriels représentatifs d'un pourcentage suffisant de la variance totale, ce qui entraîne une réduction de la dimension du nouveau système.

Plusieurs méthodes relèvent de la même démarche :

- Analyse en Composantes Principales,
- Analyse des Correspondances.

1.1. b. Méthodes de classification.

Si l'analyse factorielle a révélé l'existence de groupes et si le pourcentage de variance totale prise en compte par les trois premiers axes n'est pas significativement important, il faut approfondir l'analyse de ces données et essayer de confirmer ces résultats par des méthodes de classification ou de groupages.

Leur but est de définir une partition des éléments de l'échantillon en groupes homogènes et distincts.

On peut distinguer les méthodes hiérarchiques des méthodes non hiérarchiques.

- Classification hiérarchique :

Cette analyse permet de construire un arbre de classification (dendogramme) dont les différents niveaux représentent une partition de l'échantillon. Cet arbre sert à classer les patients et les paramètres et procède en trois étapes :

- Codification des données.
- Choix de la distance adéquate entre les éléments à classer.
- Choix d'un algorithme d'aggrégation.

- Classification non hiérarchique ou dynamique :

Elle débouche sur une partition optimale des éléments à classer. Le nombre K de groupes est fixé à l'avance, car c'est la base de la procédure, certaines classes pouvant être vides si K a été surestimé.

Procédure :

- K éléments tirés au hasard de l'échantillon et appelés 'éléments standards' constituent les noyaux des K classes.
- On calcule les distances entre chaque élément de l'échantillon et les différents éléments standards puis on alloue chaque élément au noyau le plus proche. On obtient donc K classes disjointes.
- Dans chaque classe, on prend un élément au hasard, qui constitue un nouvel élément standard. Un nouveau calcul des distances fournit une génération de K nouvelles classes disjointes.
- On itère ainsi le processus jusqu'à stabilité de la partition.

1.1. c. Méthodes décisionnelles.

L'un des paramètres joue ici un rôle spécial; c'est la variable dépendante, que l'on oppose aux autres variables qui constituent l'ensemble des variables indépendantes ou explicatives.

méthode :

- On détermine, d'après l'échantillon, si la variable dépendante est significativement expliquée par les autres.
- Si la première étape révèle une relation significative on peut prédire la valeur de la variable dépendante pour une nouvelle observation.

Parmi les principales méthodes décisionnelles applicables, selon la nature des variables (continue, discrète, binaire), citons :

- l'Analyse des Correspondances,
- l'Analyse en Composantes Principales,
- l'Analyse Discriminante Linéaire,
- l'Analyse Discriminante Quadratique.

1.1. d. Méthodologie de l'analyse des correspondances.

Pour mettre en lumière les principes énoncés, nous allons exposer la méthodologie de l'analyse des correspondances, développées par Fisher puis surtout par Benzécri.

C'est une méthode à la fois descriptive et explicative qui débouche sur des règles d'interprétations très simples.

Considérons T, la table des données à analyser,

J, le vecteur des paramètres (les colonnes de T),

I, le vecteur des observations (lignes de T),

Les données $t(i,j)$ sont d'abord traduites en fréquences $y(i,j)$ de manière à rendre les observations comparables.

$$y(i,j) = \frac{f_{ij}}{f_{i.}} \quad i \in I \text{ et } j \in J$$

$$\text{où } f_{ij} = \frac{t(i,j)}{\sum_i \sum_j t(i,j)}$$

$$f_{i.} = \sum_j f_{ij}$$

On procède ensuite au calcul de la distance CHI-CARRE entre observations :

$$d^2(i,k) = \sum_j \frac{1}{f_{.j}} \left[\frac{f_{ij}}{f_{i.}} - \frac{f_{kj}}{f_{k.}} \right]^2$$

La recherche du nouveau système de référence généré par les axes factoriels correspondant aux directions de variance maximum passe par la résolution de l'équation :

$$M V \varphi = \lambda \varphi$$

où M est la matrice des distances calculées,

V est la matrice de variance-covariance calculée à partir des $y(i,j)$,

φ et λ sont respectivement les vecteurs propres orthonormés et les valeurs propres associées de la matrice MV.

Les coordonnées des observations sur l'axe factoriel F_r

s'obtiennent par :

$$\psi_{ir} = \frac{1}{\sqrt{\lambda_r}} * \sum_j \frac{f_{ij}}{f_{i.}} \varphi_{jr}$$

où φ_{jr} représente le j-ème composant du r-ème vecteur propre.

En appliquant la même démarche au vecteur des paramètres, on obtient les formules de transition suivantes :

$$\psi_{ir} = \frac{1}{\sqrt{\lambda_r}} * \sum_j \frac{f_{ij}}{f_{i.}} \varphi_{jr}$$

$$\psi_{jr} = \frac{1}{\sqrt{\lambda_r}} * \sum_i \frac{f_{ij}}{f_{.j}} \varphi_{ir}$$

Ces formules de transition nous permettent donc de représenter les observations et les paramètres sur un seul et même plan. La proximité d'une observation et d'un paramètre sur ce plan traduit le fait que cette observation est particulièrement bien expliquée par ce paramètre.

Cela nous permet d'établir des groupes.

Toute nouvelle observation (ou nouveau paramètre) peut alors être reportée sur le plan grâce à l'application des formules de transition, et sa position est interprétée plus ou moins aisément, selon les cas.

Conclusion :

En résumé, selon la nature des données utilisées, il s'agit d'employer :

- la méthode d'analyse factorielle adéquate en vue d'obtenir une description générale des données et, si possible, une réduction de ces données,
- la méthode de classification adéquate pour confirmer un éventuel groupage de patients révélé par l'analyse factorielle,
- la méthode explicative (ou décisionnelle) adéquate, pour discriminer les groupes en obtenant une règle de prise de décision.

critique :

Les méthodes d'analyse de données multidimensionnelles n'utilisent que des variables quantitatives (résultats de dosage,...). Elles ne permettent pas de prendre directement en compte les signes cliniques, discrets et qualitatifs, qui se bornent à déterminer la présence ou l'absence d'un attribut.

D'un autre côté, elles ont la manie de classer tous les cas envisagés, si bien qu'une affection évidente se trouve subitement entachée d'un doute.

1.2. Les méthodes de discrimination non paramétrique.

principe :

A partir d'échantillons de malades de diagnostic connu, on se propose d'établir des règles de classement sans référence à des lois de probabilité. Ce type de méthode est utilisé dans des cas très particuliers.

exemple : Méthode de diagnostic par index clinique.

On recherche un certain nombre de signes considérés comme étant de poids diagnostique important. Ces signes sont affectés d'une pondération plus ou moins arbitraire.

On additionne les poids des signes dont on relève la présence pour établir la valeur d'un index.

Une étude préliminaire sur un fichier de malades cliniquement diagnostiqués permet de connaître la répartition des valeurs de l'index ainsi établi en fonction des diagnostics. Il existe parfois des zones de recouvrement où l'établissement d'un diagnostic reste douteux.

Cette méthode a été appliquée par de nombreux thyroïdologues. Il s'avère que l'index le plus simple, mis au point par R. Simonin (nous en reparlerons lors de l'élaboration des fonctions booléennes relatives à la thyroïde) en 1966 (pondération des signes = 1 si le signe est présent et = 0 sinon), est souvent le plus performant.

critique :

Cette méthode ne peut s'appliquer que dans certains cas très spécifiques, pour des pathologies où le nombre de symptômes et de diagnostics est limité. De plus, elle est mise en défaut par la présence d'un symptôme rare.

§ 2. Le diagnostic pas à pas.

2.1. Les méthodes logiques.

Ces méthodes se basent essentiellement sur la logique booléenne. Mais elles dépassent ce cadre jugé trop contraignant en s'étendant, comme nous allons successivement le montrer à la logique ternaire et aux principes régissant la démonstration automatique des théorèmes.

2.1.1. Application du calcul booléen.

Cette optique tente de classer le malade, d'abord dans un cadre nosologique assez large, puis dans des cadres de plus en plus restreints, au fil des pas diagnostiques.

Lorsque le médecin consulte le malade, il ne peut prendre en compte simultanément plus de quelques signes. La méthode logique adopte cette démarche en utilisant, à chaque itération, un nombre restreint de signes différents.

Pour mettre en lumière ces pas diagnostiques, nous allons distinguer deux phases, la première explicative, la seconde de classification.

- phase explicative :

Il s'agit de la phase la plus complexe de la démarche. En effet, il faut sélectionner, parmi tous les signes susceptibles d'être recueillis, ceux qui orientent le mieux vers tel ou tel diagnostic. Lorsque ces signes et/ou examens ont été choisis, il faut trouver leurs associations qui déterminent des groupes diagnostiques les plus précis possible.

Deux méthodes sont utilisées par la phase explicative :

- la première est basée sur le test de CHI-CARRE.

(CHU Timone de Marseille)

- la seconde s'inspire de la théorie des graphes.

(Begon, Trémolières)

- phase de classification :

Cette phase de classification est assez simple car elle se borne à l'application des fonctions logiques obtenues aux signes et/ou examens des nouveaux patients à diagnostiquer.

2.1.1. A. Phase explicative.

- a. Méthode basée sur les test de CHI-CARRE.

- a. 1. Sélection des signes.

Soit S, un ensemble de signes cliniques systématiquement recueillis par un système de questionnaire. De cet ensemble, il nous faut extraire ceux dont le poids diagnostique est le plus fort. Cette sélection s'obtient par le calcul du CHI-CARRE opéré pour chaque signe sur une table de contingence du type suivant :

	D_k	\bar{D}_k	
S_1	n_1	n_3	$n_1 + n_3$
\bar{S}_1	n_2	n_4	$n_2 + n_4$
	$n_1 + n_2$	$n_3 + n_4$	$n_1 + n_2 + n_3 + n_4$

où D_k représente le diagnostic k ,

\bar{D}_k représente l'absence du diagnostic k ,

S_1 représente la présence du signe 1,

\bar{S}_1 représente l'absence du signe 1

n_i représente le nombre de patients de diagnostic k ou de diagnostic différent de k , qui présentent ou qui ne présentent pas le symptôme 1.

La valeur du CHI-CARRE détermine ainsi le poids diagnostique de chaque signe.

Cependant, il peut arriver que deux signes sélectionnés de cette façon apportent la même information pour un diagnostic donné. Il faut alors, par un examen complémentaire, éliminer, parmi les signes qui offrent le plus grand lien avec chaque diagnostic, ceux qui s'avèrent être fortement liés entre eux.

Cela peut se faire par le calcul de la distance de Hamming (poids du vecteur obtenu par l'application de l'opérateur "ou exclusif" aux vecteurs d'observations relatifs à deux patients, ce qui détermine en fait le nombre de signes différents de l'un à l'autre).

On divise alors l'échantillon de départ en deux sous-ensembles, suite aux réponses positives ou négatives faites à certains signes. Parmi ces signes, dits signes de séparation, on recherche ceux qui minimisent la distance intra-groupe, c.à.d. la somme des distances entre un patient et les autres du même groupe, et qui maximisent la distance inter-groupes, c.à.d. la distance entre un patient et tous les patients de l'autre groupe.

a.2. Selection des associations.

Il faut insister sur le fait que, dans des associations, tous les signes importent. Il ne suffit donc pas de sélectionner uniquement les signes positifs les plus fréquents. Les signes négatifs ont un poids identique.

Sur base des dossiers, il faut rechercher les implications du type suivant :

$m_i \longrightarrow D_k$ où D_k est le k-ième groupe diagnostique au pas considéré,
 m_i est le minterm construit sur un ensemble de signes.

Definition : un minterm est une association de n signes positifs ou négatifs caractéristiques d'un patient.

Il faut ainsi trouver les mintermes qui ne correspondent qu'à un seul groupe diagnostique.

Pour les découvrir, on construit un tableau associé à chacun des diagnostics possibles. On explore le dossier médical en remplissant les différentes valeurs, positives ou négatives, associées aux différents signes retenus lors de la première étape du processus, à savoir la recherche des signes les plus discriminants.

Par diagnostic, on obtient un tableau du type suivant :

D_1	S_1	S_2	S_3	S_4	..	S_i	..
patient 1	1	0	0	1	..	1	..
patient 2	1	0	1	0	..	1	..
patient 3	1	1	0	1	..	1	..
...							
patient m

où une colonne représente les valeurs d'un signe chez les différents patients,

une ligne représente les symptômes présentés par un patient.

Pour trouver les minterms spécifiques à un diagnostic, on élimine les associations de signes, c.à.d. les combinaisons de 0 et de 1 qui sont communes à ces différents tableaux.

C'est ici que l'on se rend compte de l'utilité de l'hypothèse d'exclusivité des diagnostics; elle permet en effet d'éviter la constitution des tableaux relatifs aux combinaisons de diagnostics.

Les combinaisons qui subsistent alors dans le tableau relatif à chaque diagnostic sont donc spécifiques à ce diagnostic.

Pour les k groupes diagnostiques du pas considéré, on a ainsi sélectionné un certain nombre de minterms qui vérifient l'implication $m_i \longrightarrow D_k$. On peut alors construire, par groupe diagnostique k , une fonction F_k :

$$F_k = \sum_i m_i \quad i \in I, \text{ ensemble des minterms qui vérifient la première relation.}$$

Cette fonction adopte une forme disjonctive.

L'ensemble des associations de signes identifiées peuvent se représenter sous forme d'un tableau :

SIGNES DISCRIMINANTS	S_1				
	S_2	m_1	m_2	\dots	m_i
	S_3				
	\dots				
	S_m				
DIAGNOSTICS POSSIBLES	D_1	X	X		
	D_2			X	
	\dots				
	D_n				x

exemple tiré de la pathologie thyroïdienne :

On détermine ceux qui parmi les minterms vérifient par exemple les dossiers déjà diagnostiqués hypothyroïdiens et parmi ceux-là choisir ceux

qui vérifient le moins possible les dossiers hyperthyroïdiens, normaux ou de thyroïse. C'est ainsi qu'ont été élaborées les fonctions d'aide au diagnostic thyroïdien par simple addition des miniterms sélectionnés. Les résultats obtenus seront développés dans un chapitre ultérieur.

b. Méthode basée sur la théorie des graphes.

Comme pour la méthode utilisée au C.H.U. Timone de Marseille, elle a pour objectif de proposer des signes et des associations de signes et/ou d'examens impliquant des groupes de diagnostics.

Nous avons déjà soulevé au point précédent, le problème du choix des signes. Dans cette optique, nous traiterons du choix sous critère. En effet, chaque examen s'accommode d'inconvénients dont il nous faut tenir compte, inconvénients de coût, de durée, etc....

Nous posons la même hypothèse d'exclusivité des diagnostics.

Nous développerons un exemple théorique qui met en jeu cinq diagnostics, six examens et deux critères.

Soient E_j , les examens considérés,

Soient D_i , les cinq diagnostics possibles.

CRITERES		TABLE de DIAGNOSTICS					
COUT (en B)	TEMPS (en minutes)	EXAMENS	MALADIES				
40	4	E_1	0	0	0	0	1
15	7	E_2	1	0	0	0	0
50	4	E_3	1	1	1	0	1
16	5	E_4	0	0	0	1	1
75	20	E_5	0	1	0	0	1
8	5	E_6	1	0	1	0	0
			D_1	D_2	D_3	D_4	D_5

Sur ce tableau, une ligne montre le profil d'un examen pour les différents diagnostics tandis qu'une colonne traduit le profil d'un diagnostic au travers des différents examens.

- principe :

La démarche diagnostique proposée par Begon-Trémolières consiste à déterminer tous les sous-ensembles d'examens permettant de distinguer les différents diagnostics (ces ensembles étant appelés ensembles filtrants), et à choisir en fonction de certains critères, l'un de ces sous-ensembles.

- méthode :

Dans un premier temps, on va chercher tous les sous-ensembles d'examens qui différencient à la fois D_1 de D_2 et D_1 de D_3 et D_4 de D_5 , ce qui donne le tableau suivant, appelé tableau des ET :

POUR DISTINGUER	IL FAUT FAIRE LES EXAMENS	
ET (D^1 de D^2)	E_2 ou E_5 ou E_6) ET	L_1
ET (- de D^3)	E_2) ET	L_2
- de D^4	E_2 ou E_3 ou E_4 ou E_6) ET	L_3
ET (- de D^5)	E_1 ou E_2 ou E_4 ou E_5 ou E_6)	L_4
. D^2 de D^3	E_5 ou E_6 .	L_5
. - de D^4	E_3 ou E_4 ou E_5 .	L_6
. - de D^5	E_1 ou E_4 .	L_7
. D^3 de D^4	E_3 ou E_4 ou E_6) ET	L_8
ET (- de D^5)	E_1 ou E_4 ou E_5 ou E_6) ET	L_9
ET (D^4 de D^5)	E_1 ou E_3 ou E_5)	L_{10}

Dans un deuxième temps, nous rechercherons les sous-ensembles filtrants.

b.1. Méthode par énumération.

Il est aisé de voir que les sous-ensembles d'examens d'un ou de deux éléments sont insuffisants pour aboutir à un diagnostic valable. En effet, quatre combinaisons sont réalisables avec deux examens pour cinq diagnostics possibles. Au minimum, trois examens seront donc nécessaires pour discriminer les diagnostics.

D'une façon générale, le nombre minimum d'examens nécessaires pour différencier n diagnostics doit être tel que

$$2^p \gg n$$

La méthode par énumération est très simple. Elle consiste à énumérer tous les sous-ensembles d'examens comprenant au minimum p éléments et de voir s'ils sont filtrants ou non. Rappelons qu'un sous-ensemble d'examens est filtrant s'il est discriminant pour chaque couple de diagnostics, c.à.d. lorsqu'il y a au moins un élément commun avec chacun des sous-ensembles du tableau des ET.

Le tableau suivant présente les sous-ensembles filtrants obtenus :

N°	SOUS-ENSEMBLES FILTRANTS	N°	SOUS-ENSEMBLES FILTRANTS
1	* E ₂ E ₄ E ₅	9	E ₂ E ₄ E ₅ E ₆
2	* E ₁ E ₂ E ₃ E ₅	10	E ₁ E ₂ E ₃ E ₄ E ₅
3	* E ₁ E ₂ E ₃ E ₆	11	E ₁ E ₂ E ₃ E ₄ E ₆
4	E ₁ E ₂ E ₄ E ₅	12	E ₁ E ₂ E ₃ E ₅ E ₆
5	* E ₁ E ₂ E ₄ E ₆	13	E ₁ E ₂ E ₄ E ₅ E ₆
6	* E ₁ E ₂ E ₅ E ₆	14	E ₂ E ₃ E ₄ E ₅ E ₆
7	E ₂ E ₃ E ₄ E ₅	15	E ₁ E ₂ E ₃ E ₄ E ₅ E ₆
8	* E ₂ E ₃ E ₄ E ₆		

Les sous-ensembles marqués '*' sont dits "de base", ce qui signifie qu'aucun autre sous-ensemble filtrant ne leur est inclus.

Pour tenir compte des critères, il suffit de pondérer tous les sous-ensembles filtrants de base par la somme des coûts ou des durées et de choisir celui qui optimise le critère considéré.

Cependant, lorsque le nombre d'examens et de diagnostics augmente, cette méthode devient lente et fastidieuse. C'est pourquoi les méthodes graphiques ont été introduites.

b.2 Méthode graphique d'exploration directe.

Cette méthode nécessite également l'élaboration préliminaire du tableau des ET, que l'on restructure en ordonnant ses lignes par nombre croissant d'éléments. Cette restructuration accélérera l'obtention des sous-ensembles filtrants.

Une deuxième manipulation de ce tableau consiste à restructurer chaque ligne du tableau ordonné en plaçant les examens dans l'ordre résultant du choix d'un critère.

Dans le tableau présenté ci-dessous, nous n'avons tenu compte que du critère coût, l'ordre utilisé étant celui des coûts croissants.

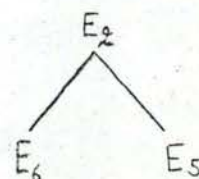
T_1	E_2
T_2	E_6 OU E_5
T_3	E_4 OU E_1
T_4	E_6 OU E_2 OU E_5
T_5	E_4 OU E_3 OU E_5
T_6	E_6 OU E_4 OU E_3
T_7	E_1 OU E_3 OU E_5
T_8	E_6 OU E_2 OU E_4 OU E_3
T_9	E_6 OU E_4 OU E_1 OU E_5
T_{10}	E_6 OU E_2 OU E_4 OU E_1 OU E_5

Ce tableau sert de base à la construction d'une arborescence où l'on adopte la règle "toujours à gauche".

Suivons les différentes étapes de la construction de cette arborescence.

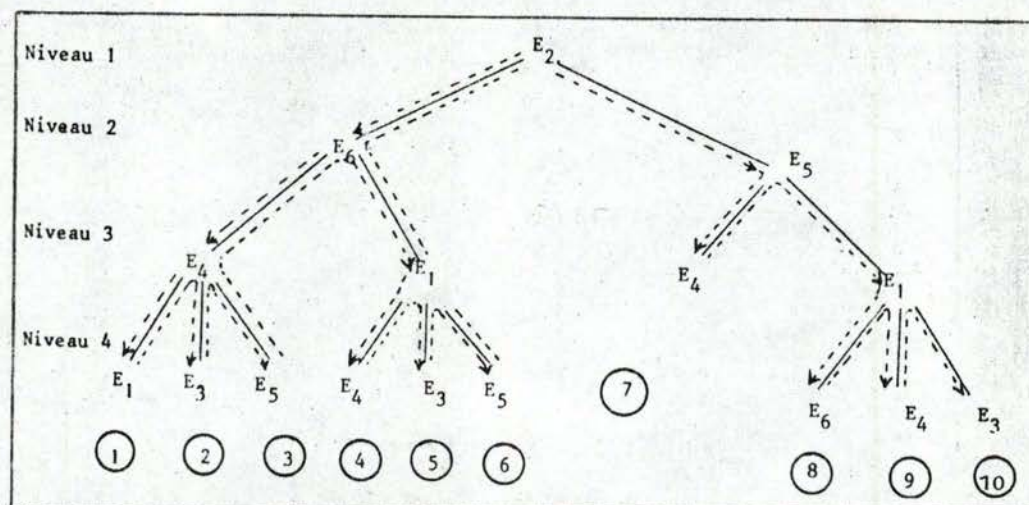
- étape 1 : On met la première ligne du dernier tableau présenté à la tête de l'arborescence. La racine est donc, dans notre exemple, E_2 .
- étape 2 : On construit le second niveau avec les éléments de la deuxième ligne du même tableau à la condition qu'aucun élément de cette ligne ne figure au niveau supérieur.

On obtient pour notre exemple le graphe suivant :



- étape 1 : On construit le niveau suivant avec les éléments de la i -ième ligne du tableau. Ces éléments sont ajoutés à chaque branche de l'arborescence à la condition qu'aucun élément de la i -ième ligne ne figure pas dans un niveau supérieur.

Notre exemple aboutit à l'élaboration du graphe suivant :



Cette méthode permet d'exhiber rapidement une ou plusieurs chaînes filtrantes. En effet, une fois l'arborescence construite, la lecture "toujours à gauche" nous permet d'obtenir un dernier tableau qui reprend, en regard des sous-ensembles filtrants obtenus, le poids du critère coût.

EXPLORATION DIRECTE AVEC LE CRITERE COUT		
Ordre d'obtention des chaînes	sous-ensembles filtrants	coût (en B)
1	E_1, E_2, E_4, E_6	79
2	E_2, E_3, E_4, E_6	89
3	E_2, E_4, E_5, E_6	114
4	<i>idem</i> (1)	
5	E_1, E_2, E_3, E_6	113 *
6	E_1, E_2, E_5, E_6	138
7	E_2, E_4, E_5	106 *
8	<i>idem</i> (6)	
9	E_1, E_2, E_4, E_5	146
10	E_1, E_2, E_3, E_5	180

Cette méthode de construction est une adaptation de la procédure LIFO (Last In First Out) de la méthode du "Branch and Bound" ou procédure S.E.P. (procédure par Séparation et Evaluation Progressive). Le dernier tableau montre que les premières chaînes obtenues sont celles qui minimisent le facteur coût. On constate cependant que l'ordre d'obtention de ces chaînes et la progression de leurs poids ne sont pas strictement parallèles (voir *).

Si, dans cet exemple, les premières chaînes exhibées sont celles qui coûtent le moins cher, il ne peut en être toujours de même. Cette méthode d'exploration directe est heuristique.

En résumé, elle guide le choix des examens complémentaires. Elle pourrait se coupler avec la méthode utilisée au C.H.U. Timone de Marseille en ce qui concerne l'élaboration des fonctions employées lors du diagnostic proprement dit.

2.1.1.B. Phase de classification.

C'est la phase du diagnostic.

Si le malade présente une symptomatologie qui donne à la fonction F_k la valeur vraie, alors on est sûr qu'il a la maladie k . Par contre, si toutes les fonctions F_k donnent la valeur fausse pour un pas donné, il est impossible de classer le malade car les informations connues à son sujet ne sont provisoirement pas suffisantes.

Il existe donc, non pas une branche d'erreur, mais une branche de rejet. Cette branche de rejet sera ensuite traitée de la même manière que les autres sous-branches d'affinement.

On pourrait définir cette méthode diagnostique pas à pas: comme un enchaînement de décisions élémentaires auxquelles est soumis le patient et dont les éléments sont les fonctions logiques précédemment définies.

Comme on le voit, cette méthode suit exactement la démarche du médecin qui prend en compte, à chaque pas, une série de nouveaux signes. A chaque itération, on incorpore ces signes dans les différentes fonctions logiques pour affiner le diagnostic.

2.1.2. Extension à l'algèbre ternaire.

- principe :

Toute variable ne peut prendre que trois valeurs (1,0,-1) selon respectivement la présence, l'incertitude ou l'absence du signe dans la symptomatologie du patient considéré. Les relations entre symptômes et maladies s'expriment sous forme de relations logiques et peuvent se représenter sur une table de décision, après diverses simplifications analogues à celles opérées dans les autres méthodes logiques.

L'extension de l'algèbre de Boole à une algèbre à trois états a cependant nécessité la création d'opérateurs supplémentaires.

Voici une des façons de considérer le problème, celle adoptée par une équipe de recherches du Pr. Grémy (groupe de recherches U88-INSERM-PARIS).

- + 1 correspond à la présence.
- 0 correspond à l'incertitude.
- 1 correspond à l'absence.

Le OU binaire devient le maximum.

Le ET binaire devient le minimum.

Le Complément devient la négation algébrique.

% signifie que l'incertitude devient absence.

@ correspond à la rotation positive.

Les opérateurs ternaires utilisés apparaissent sur les deux tables de vérité suivantes :

x	\bar{x}	@ x	% x
1	-1	-1	1
0	0	1	-1
-1	1	0	-1

x	y	x + y	x . y
0	0	0	0
0	1	1	0
0	-1	0	-1
1	0	1	0
1	1	1	1
1	-1	1	-1
-1	0	0	-1
-1	1	1	-1
-1	-1	-1	-1

2.1.3. La démonstration automatique de théorèmes.

- principe :

Cette méthode se fonde également sur les déductions logiques et elle peut intervenir à tous les pas diagnostiques.

La connaissance médicale relative à chaque diagnostic est mise sous forme d'axiomes. Ceux-ci constituent ce que l'on appelle, en logique du premier ordre, une théorie. L'association des signes présentés par le patient est considérée comme une formule de la théorie. Si cette formule peut se ramener à un des axiomes, alors le patient est affecté d'une maladie, celle dont il est question dans l'axiome.

- critique des méthodes logiques.

Les méthodes logiques ne tiennent pas compte des arguments de fréquence. Elles se contentent d'exprimer la présence ou l'absence d'un symptôme sans tenir compte des fréquences d'apparition observées de cette existence ou de cette absence.

Elles peuvent néanmoins être le mieux adaptée pour certains types de problèmes. On le verra lors de l'étude de résultat relative à la pathologie thyroïdienne.

2.2 La méthode probabiliste (théorème de Bayes).

- hypothèses de base :

Exclusivité et exhaustivité des diagnostics.

Pour être applicable, il faut qu'un patient donné ne soit atteint que d'une et une seule des maladies possibles, l'éventail de ces maladies devant être complet.

Indépendance des signes recherchés.

La réponse à un signe ne peut en aucun cas dépendre des réponses obtenues aux signes précédemment recherchés. Pareille hypothèse est cependant bien difficile à respecter au plan pratique. Aussi, certains auteurs ont essayé d'y porter remède en introduisant dans le processus bayésien la dépendance des signes,

- soit en cherchant à regrouper les signes deux à deux, en calculant le CHI-CARRE de toutes les paires possibles (Nugent),
- soit en substituant à l'hypothèse d'indépendance totale celle, non moins contraignante, de relations linéaires entre signes (Brunk, Lehr),
- soit en regroupant les signes exclusifs de sorte qu'un seul signe de chaque groupe puisse être pris en compte dans le calcul du diagnostic (Warner, Toronto et Veasy),
- soit encore en établissant une formule qui tienne compte des corrélations de tout ordre entre les signes (Bahadur) ou en construisant un modèle original adapté aux signes dépendants.

Cependant, ces diverses tentatives ne restent applicables que pour un nombre de signes et de diagnostics relativement faible .

Ce problème de l'indépendance des signes est volontiers exagéré ou minimisé. A ce sujet, introduisons ici trois remarques :

- L'hypothèse d'indépendance des signes n'est pas absolument nécessaire à l'utilisation d'un processus bayésien qui, sur le plan théorique, peut tout à fait s'en passer. Seules des considérations pratiques font que, le plus souvent, les expériences réalisées doivent assumer une telle hypothèse.

- Cette hypothèse est relative à chacun des diagnostics considérés. Cela revient à dire que la probabilité d'existence d'un signe quelconque, si tel diagnostic est présent, est la même si tel autre signe existe ou n'existe pas, ce qui ne traduit pas l'indépendance absolue entre signes :

$$P(S_i/S_j) = P(S_i)$$

Ce fait met en lumière l'un des intérêts de la première hypothèse, à savoir l'exclusivité des diagnostics, qui permet plus facilement d'analyser et éventuellement d'accepter certaines hypothèses d'indépendance.

- Bien souvent, les auteurs acceptent l'hypothèse de "pseudo-indépendance" qui élimine uniquement les relations les plus fortes entre signes.

- paramétrage :

La méthode probabiliste consiste à traduire un fichier de données, celui des dossiers médicaux, en probabilités, sur lesquelles se fonde l'application du théorème de Bayes.

Définissons les variables :

- ND_j , le nombre de malades de diagnostic D_j ,
- N_i , le nombre de malades qui présentent le signe S_i ,
- N_i^+ , le nombre de malades qui ont le symptôme S_i positif,
- N_i^- , le nombre de malades qui ont le symptôme S_i négatif,
- N_{ij} , le nombre de malades de diagnostic D_j qui présentent le signe S_i ,
- N_{ij}^+ , le nombre de malades de diagnostic D_j qui présente le symptôme S_i positif,
- N_{ij}^- , le nombre de malades de diagnostic D_j qui présentent le symptôme S_i négatif.

L'on peut donc estimer les paramètres suivants sur base d'un échantillon de malades de diagnostic confirmé.

- probabilité de chaque diagnostic.

$$P(D_j) = \frac{ND_j}{\sum_k ND_k}$$

où $P(D_j)$ représente la probabilité à priori du diagnostic D_j .

- probabilité de chaque signe.

$$P(S_i) = \frac{N_i^+}{N_i}$$

$$Q(S_i) = \frac{N_i^-}{N_i} = 1 - P(S_i) \text{ où } P(S_i) \text{ représente la probabilité à priori que le symptôme } S_i \text{ apparaisse de manière positive.}$$

- probabilité de chaque signe relative à chaque diagnostic.

$$P(S_i/D_j) = P(S_i^+ / D_j) = \frac{N_{ij}^+}{N_{ij}}$$

$$Q(S_i/D_j) = \frac{N_{ij}^-}{N_{ij}}$$

où $P(S_i/D_j)$ représente la probabilité conditionnelle de S_i par rapport à D_j .

Comme nous le verrons plus loin, l'estimation de ces différents paramètres pose bien souvent de nombreux problèmes pratiques.

- calcul du diagnostic :

La connaissance d'un signe S_i permet d'ajuster les probabilités a priori des différents diagnostics grâce à l'application pure et simple du théorème de Bayes :

$$P(D_j/S_i) = \frac{P^o(D_j) * P(S_i/D_j)}{\sum_k P^o(D_k) * P(S_i/D_k)}$$

où $P^o(D_j)$ représente la probabilité du diagnostic D_j avant la prise en compte de l'observation du signe S_i . Initialement, lorsqu'aucun signe n'est encore considéré, $P^o(D_j)$ représente la probabilité a priori du diagnostic D_j .

L'hypothèse d'indépendance des symptômes permet d'affirmer par exemple que,

$$P(D_j/S_1^+, S_2^-, S_3^+) = \frac{P^o(D_j) * P(S_1/D_j) * Q(S_2/D_j) * P(S_3/D_j)}{\sum_k P^o(D_k) * P(S_1/D_k) * Q(S_2/D_k) * P(S_3/D_k)}$$

De plus, l'absence d'un symptôme peut présenter autant d'intérêt que la présence d'un autre et le théorème de Bayes permet précisément d'en tenir compte.

La prise en considération des signes observés chez le patient modifie la valeur du vecteur des probabilités des diagnostics en probabilités a posteriori. Pour interpréter ces valeurs de probabilités, nous proposons trois critères différents :

- considérer l'ordre (probabilité maximum = diagnostic réel),
- établir un seuil de décidabilité qui fixe un diagnostic en cas de dépassement de ce seuil, qui propose des examens complémentaires ou qui laisse planer le doute.
- considérer la cardinalité des probabilités des diagnostics, dans l'espoir d'y trouver des différences significatives entre les premiers classés.

- critique :

L'estimation des probabilités a priori et des probabilités conditionnelles nécessite des statistiques convenablement rassemblées et mises à jour, ce qui n'est pas toujours le cas dans certains milieux médicaux.

Lorsque l'échantillon se révèle trop petit, on fait appel à des probabilités "subjectives" déterminées par des médecins spécialistes et progressivement revues et affinées.

L'expérience a montré que ces évaluations cernaient la réalité avec beaucoup de précision, mis à part quelques cas particuliers.

L'indépendance des symptômes entrant dans les différents profils des patients n'est pas toujours vérifiée, loin s'en faut. Cependant, cette lacune n'a pas nui à l'efficacité pratique des différents programmes d'aide au diagnostic actuellement mis au point.

Ainsi que nous l'avons vu, certains auteurs ont essayé d'y remédier sans obtenir pour autant des résultats sensiblement meilleurs.

L'existence de signes ou de syndromes qui excluent ou qui identifient définitivement un diagnostic ou un groupe de diagnostics n'est pas considérée.

Enfin, pour l'élaboration du diagnostic, seules les données qualitatives peuvent être directement utilisées si bien que les données quantitatives, réduites à des valeurs qualitatives (intervalles de valeurs, groupages, ...) perdent un peu de leur valeur informationnelle.

2.3. La théorie des sous-ensembles flous.

- principe :

Tout développement mathématique doit suivre la perception humaine. S'il est impossible d'affecter des valeurs précises dans notre pensée, notre perception doit toujours être exacte et fidèle. Faute pour l'imprécis mais rejetant l'inexactitude, telle apparaît la théorie des sous-ensembles flous. C'est peut-être parce que tant de gens ont confondu l'imprécis et l'inexact qu'une injuste méfiance a prévalu pendant plusieurs années vis-à-vis de cette théorie.

La théorie des sous-ensembles flous est une théorie des connaissances imprécises et qui ne peut accepter le raisonnement inexact; elle se propose de nous faire utiliser des modèles où l'on peut accepter des données, des relations, des procédures, des structures qui sont floues, c.à.d. qui sont perçues et utilisées avec leur imprécision naturelle, mais dans le but d'obtenir des décisions qui concordent au mieux avec les intentions des décideurs. Cette théorie est fondée sur la notion d'appartenance.

Mais on ne se limite pas aux deux phénomènes classiques d'appartenance ou de non-appartenance.

Si l'on se place dans un contexte médical avec un ensemble de patients, peut-on dire d'une personne : elle appartient ou elle n'appartient pas à l'ensemble de maladies ?

Non, car les limites d'une telle affirmation sont floues.

C'est pour cette raison que l'on a introduit la notion de "degré d'appartenance".

On voit très bien l'utilité de cet outil lorsque l'on travaille sur des critères tels que :

- une personne a-t-elle de l'humour ?
- une personne est-elle irritable, jolie ou intelligente ?
- une personne est-elle atteinte de telle maladie mentale ?

Ainsi, les réponses à de telles questions et l'établissement de pareils diagnostics entraînent des constructions abstraites qui constituent les sous-ensembles flous, c.à.d. des classes aux limites imprécises dans lesquelles la transition d'appartenance est plus graduelle que brusque.

Cette théorie développée par le Pr. L.A. Zadeh de l'université de Berkeley a été introduite en France par le Pr. A. Kaufmann.

Un sous-ensemble flou se définit par la liste des différents éléments auxquels est attribué un degré d'appartenance à un ensemble de référence et non, comme pour les ensembles traditionnels, par l'énumération des objets qui le composent.

Dès à présent, nous pouvons dire que l'idée qui est à la base de cette théorie s'adapte bien au domaine médical. En effet, le médecin est souvent confronté à des données, des appréciations imparfaitement définies et il est amené à interpréter ces informations en tant que telles dans leur propre domaine non mesurable. De plus, il existe souvent des trous d'information.

Cette théorie permet aussi de prendre en compte l'importance qu'il convient d'attribuer à l'absence d'un symptôme qui, pour diverses raisons, n'a pas été recherché.

Une équipe de Marseille, sous l'impulsion de R. Sambuc a généralisé la notion de sous-ensemble flou en introduisant le concept de sous-ensemble $\bar{\Phi}$ -flou qui permet de traduire à la fois la notion de flou qui caractérise l'appartenance pondérée à un ensemble et la notion d'indétermination qui peut s'attacher à l'évaluation même de cette appartenance.

°

° °

CHAPITRE II

APPLICATIONS ET ÉTUDES DE RÉSULTATS

Les différentes applications que nous verrons lors de ce chapitre ont été réalisées par le laboratoire de biomathématiques, statistiques, informatique médicale du Professeur M. Roux du C.H.U. Timone de Marseille.

Ces différentes études se rapportent toutes à la Pathologie thyroïdienne.

Comme base de travail, l'équipe de Marseille s'est appuyée sur son dossier médical : SYCVAR que nous verrons dans la deuxième partie.

Les Méthodes appliquées et qui seront exposées dans ce chapitre sont les suivantes :

1. Méthode basée sur la logique booléenne.
2. Méthode basée sur la logique floue.
3. Méthode de type multidimensionnelle.
4. Méthode par pondération des caractères qualitatifs.

Nous tirerons quelques conclusions sur ces méthodes en un point 5.

5. Conclusions sur ces différentes applications.

1. Méthode basée sur la logique booléenne

Cette application thyroïdienne a exactement suivi le processus proposé par le C.H.U. Timone de Marseille et étudié dans le chapitre: Méthodologie.

Après la sélection des signes les plus discriminants et leurs associations, l'équipe de Marseille a obtenu quatre fonctions représentatives des quatre groupes diagnostiques relatifs à la première étape de la démarche pas à pas.

Il s'agit des groupes diagnostiques suivants :

- Hypothyroïde	représenté	par	la	fonction	L 1
- normal	"	"	"	"	L 2
- Hyperthyroïde	"	"	"	"	L 3
- Thyréose	"	"	"	"	L 4

Ces quatre fonctions ont été obtenues sur l'analyse du dossier SYCVAR et sont les suivantes :

L 1 : (Frilosité ou asthénie après effort) et (peau sèche et squameuse ou augmentation de poids) et (taux de fixation 6e heure diminué ou taux de fixation 24e heure diminué) et Blanquet Meyniel III diminué et (réflexogramme VP 200 ms ou voix rauque et nasonnée)..

L 2 : (non Blanquet Meyniel III augmenté) et (non Blanquet Meyniel III diminué) et (non taux de fixation 24e heure augmenté) et (non taux de fixation 24e heure diminué).

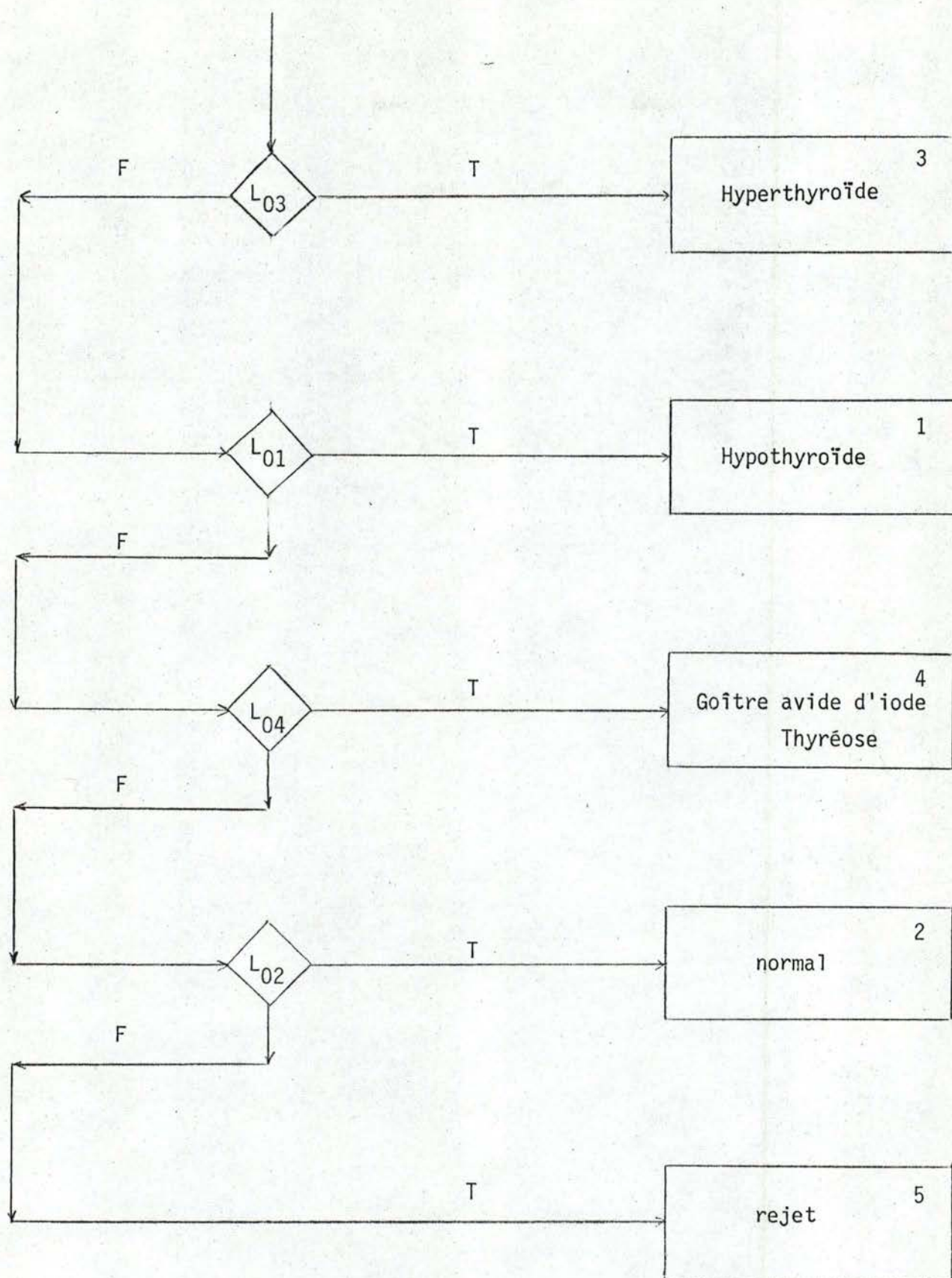
L 3 : (Tachycardie ou amaigrissement) et (thermophobie ou soif) et (taux de fixation 6e heure augmenté ou taux de fixation 24e heure augmenté) et Blanquet Meyniel III augmenté.

L 4 : Taux de fixation 24e heure augmenté et taux de fixation 48e heure augmenté et Blanquet Meyniel III augmenté et (non tachycardie ou non amaigrissement) et (non thermophobie ou non soif).

L'application de ces fonctions se fait selon l'organigramme ci-après.
Ce qui constitue le schéma du premier pas diagnostique.

Etape d'orientation

Etape d'affinement



Le qui est surtout intéressant, c'est de montrer la pertinence des résultats obtenus grâce à ces fonctions lors de la première étape diagnostique.

Le test s'est effectué sur mille nouveaux patients. 324 d'entre eux offraient des médications perturbatrices ou des diagnostics n'ayant aucun rapport avec la pathologie thyroïdienne. Les fonctions vues ne tiennent pas compte de ces cas. Il restait 676 dossiers qui allaient être soumis aux quatre fonctions logiques que nous avons vues.

Résultats.

N° diagnostic effectif N° fonc- tions vérifiées	1	2	3	4	5
	25	455	134	47	TOTAL
L_1	15	0	0	0	15
L_2	3	294	0	6	303
L_3	0	0	94	0	94
L_4	0	9	1	33	43
Tableau 1					455

On constate que 455 dossiers sur 676 ont été classés, ce qui nous donne donc 67 % de patients classés (d'où également 33 % de rejet c.a.d. manque d'informations pour décider; ces 33 % pourront être intégrés dans les différents diagnostics lors des étapes d'affinement). De plus, on désirerait savoir si un malade qui répond à une des fonctions L_i est classé dans un des groupes avec quasi-certitude ! On peut estimer cette quasi-certitude en employant le calcul des probabilités.

On désire estimer la probabilité qu'un malade vérifiant une des quatre fonctions logiques se trouve réellement dans le groupe diagnostique correspondant ou encore exprimé sous forme mathématique. $Pr(D_i/L_i)$ ou $i = 1, 4$

De manière plus générale, nous avons élaboré un tableau où sont reprises toutes les probabilités conditionnelles $Pr(D_i/L_h)$ où $i = 1, 4$ et où $h = 1, 4$

L_1	L_2	L_3	L_4
$P(D_1/L_1) = 1.0$	$P(D_1/L_2) = 0.01$	$P(D_1/L_3) = 0.0$	$P(D_1/L_4) = 0.0$
$P(D_2/L_1) = 0.0$	$P(D_2/L_2) = 0.97$	$P(D_2/L_3) = 0.0$	$P(D_2/L_4) = 0.21$
$P(D_3/L_1) = 0.0$	$P(D_3/L_2) = 0.0$	$P(D_3/L_3) = 1.0$	$P(D_3/L_4) = 0.02$
$P(D_4/L_1) = 0.0$	$P(D_4/L_2) = 0.02$	$P(D_4/L_3) = 0.0$	$P(D_4/L_4) = 0.77$
$P(L_1/D_1) = 0.6$	$P(L_2/D_2) = 0.65$	$P(L_3/D_3) = 0.7$	$P(L_4/D_4) = 0.7$
Soit 40% de sujets hypothyroïdiens non retenus	Soit 35% de sujets fonctionnellement normaux non retenus	Soit 30% de sujets hyperthyroïdiens non retenus	Soit 30% de sujets thyroésiques non retenus

tableau 2

En étudiant ce tableau, on remarque que lorsqu'un patient répond à une des fonctions, son diagnostic est quasi-certain. On le constate en examinant les probabilités $Pr(D_i/L_i)$ du tableau. Ainsi, la probabilité d'avoir le diagnostic D_1 si L_1 est vraie est de 100 %, la probabilité d'avoir le diagnostic D_2 si L_2 est vraie est de 97 % et la probabilité d'avoir le diagnostic D_3 si L_3 est vraie est également de 100 %.

On remarque néanmoins qu'un certain problème existe pour la fonction L_4 . En effet, il y a une probabilité de 21 % qu'un patient vérifiant L_4 soit en fait normal. Ceci s'explique du fait qu'il n'existe pas de frontière clinique nette entre les thyroésoses et l'état normal.

On constate également que le nombre de sujets de chaque groupe non retenus reste assez élevé : de 30 à 40 % suivant les groupes.

On remarque également sur le tableau 1 que certaines classifications sont mauvaises. Cela ne doit pas être imputé à des erreurs diagnostiques mais cela vient du fait que certains sujets sont à la limite entre deux groupes diagnostiques.

Ce que nous venons de voir pour le premier pas diagnostique est parfaitement valable pour le deuxième pas..

Ce deuxième pas permet d'ailleurs de remédier à ces mauvaises classifications car les organigrammes d'affinement élaborés par les cliniciens permettent de classer dans des sous-groupes diagnostiques identiques des sujets provenant de deux groupes diagnostiques différents mais dont la symptomatologie était proche.

Ce deuxième pas permet également d'éliminer certains "rejets" engendrés lors de la première étape.

On pourrait croire que, une fois les fonctions logiques trouvées, le travail des cliniciens tendra à s'étioler. Or il n'en est rien car il est toujours possible d'augmenter la sélectivité des fonctions L initialement adoptées. Cette sélectivité étant commandée par le choix des minterms qui la constitue, le clinicien travaille toujours au test de nouveaux signes, de nouveaux examens médicaux augmentant la pertinence des résultats.

Ainsi, de nouveaux tests à très fort "poids diagnostique" ont été découverts et utilisés pour la pathologie thyroïdienne : le T3 Test, l'index de Thyroxine libre (iTL) et le dosage radio-immunologique.

On a également introduit dans les fonctions diagnostiques le concept d'index clinique. La méthode des index cliniques est relativement simple et peut s'appliquer à certains types d'affectation bien déterminés : on recherche un certain nombre de signes considérés comme de grande valeur diagnostique que l'on affecte d'une pondération plus ou moins arbitraire. La valeur de l'index est obtenue en additionnant les poids des différents signes relevés chez le malade. Certains diagnostics peuvent être caractérisés par des valeurs particulières de l'index.

L'index du docteur Simonin a été intégré dans les fonctions de diagnostic thyroïdien. Comme tous les index cliniques, les signes choisis (ici au nombre de 8) font l'objet d'une pondération. La pondération du docteur Simonin est simple : 1 si le signe est présent
0 si le signe est absent.

L'étude montre que l'on ne trouve pratiquement plus que des malades hyperthyroïdiens parmi ceux qui ont un index supérieur ou égal à 5.

Prenons un exemple de fonction modifiée par les nouveaux éléments introduits : soit Lo3 la nouvelle fonction relative à l'hyperthyroïde

Lo3 : L3 ou (index Simonin \geq 5 et T3 - Test $>$ 35 %)
ou valeur augmentée de la radio-immuno
ou ITL augmenté.

Nous allons de nouveau montrer l'accroissement des performances par quelques chiffres.

Diagnostiques réels \ Fonctions	L ₁	L ₂	L ₃	Aucune fonction satisfaite
Hypothyroïde	51	4	0	2
Normaux	4	833	3	72
Hyperthyroïde	0	1	106	10

tableau 3

On remarque qu'il demeure moins de classifications anormales. De plus, 89 % des hypothyroïdiens sont bien classés, 91 % des sujets ne présentant pas d'anomalies thyroïdiennes et 91 % des hyperthyroïdiens le sont également.

Soit globalement 90 % des sujets.

Il n'est évidemment pas à exclure que d'autres modifications pourront encore être apportées et que la pertinence des résultats en sera une fois de plus accrue.

Nous montrons en annexe 2 le type de résultat obtenu par cette méthode avec un résumé du dossier du patient et le diagnostic qui l'accompagne.

2. Méthode basée sur la logique floue

Cette méthode est calquée sur la méthode booléenne. On travaillera à partir des fonctions booléennes obtenues au point précédent. On remarque donc que l'on ne fait pas intervenir la notion du flou au niveau de la sélection des signes discriminants et de la création des associations. Le flou intervient uniquement lors de la prise en considération des signes des patients et à la transformation des fonctions booléennes connues.

2.1. Introduction aux ensembles Φ -flous.

Afin d'aboutir rapidement aux notions d'ensembles Φ -flous, nous allons rappeler quelques points importants relatifs aux ensembles vulgaires et flous. Nous nous limiterons surtout à l'aspect intuitif de ces concepts.

2.1.1. Les Ensembles vulgaires ou traditionnels.

La théorie des ensembles traditionnels est fondée sur la notion d'appartenance.

Cette fonction d'appartenance prend uniquement ses valeurs dans l'ensemble $\{0,1\}$. Il s'agit donc d'une fonction à deux valeurs.

Un ensemble traditionnel se représente notamment par l'énumération de ses composants.

Exemple : soit un ensemble A composé de trois éléments a,b,c.

Cet ensemble se note $A = \{a,b,c\}$

on aurait pu le représenter par un ensemble de couples où le premier élément serait le composant considéré et le deuxième la valeur de la fonction d'appartenance. Ainsi A pourrait se noter $\{(a,1), (b,1), (c,1), (d,0), (e,0)\}$ en considérant que d et e n'appartiennent pas à A

La logique booléenne s'inspire des mêmes notions c.à.d. qu'elle possède deux valeurs de vérité VRAI et FAUX.

Exemple : A la question : "Cette fille est-elle jolie", la logique booléenne ne peut répondre que par oui ou par non. Ce qui est tout de même assez limité pour exprimer un fait aussi subjectif.

Les variables booléennes sont les valeurs que peut prendre la fonction d'appartenance c.à.d. les valeurs 0 et 1.

2.1.2. Les Sous-Ensembles flous

La théorie des sous-ensembles flous est également basée sur la notion d'appartenance. Mais la fonction d'appartenance considérée ici ne prend plus ses valeurs dans l'ensemble $\{0,1\}$ mais sur tout l'intervalle $[0,1]$. On introduit ainsi toute une gradation dans l'appartenance d'un élément.

Remarque : on parle toujours de sous-ensembles flous car on travaille toujours par rapport à un référentiel qui est un ensemble traditionnel.

On représente ces sous-ensembles flous par un ensemble de couples où le premier élément est un composant du référentiel considéré et où le deuxième élément est son degré d'appartenance par rapport au sous-ensemble flou considéré.

Exemple 1

Soit le référentiel E défini de la façon suivante :

$$E = \{a, b, c, d\}$$

Soit le sous-ensemble flou A défini de la façon suivante :

$$A = \{(a, 0.1); (b, 0.5); (c, 1); (d, 0)\}$$

La logique floue s'inspire de ces notions c'est-à-dire qu'elle possède ses valeurs de vérité étalées sur tout l'intervalle $[0,1]$

Cette logique s'adapte bien aux imprécisions de notre pensée, tout en rejetant l'inexactitude.

Exemple 2

A la question : "cette fille est-elle jolie ?", la logique floue permet de graduer notre réponse en 0 et 1. Ceci permet de mieux rendre compte de notre pensée. En effet, à des questions aussi subjectives, on répond rarement par oui ou non mais plutôt par des réponses telles que : assez jolie, très jolie ...

La logique floue nous permet ce genre de réponses,
 ainsi 0.9 : fille très jolie
 0.6 : fille assez jolie
 0.1 : fille laide.

Les variables floues prennent les valeurs que peut prendre la fonction d'appartenance floue c'est-à-dire des valeurs de l'intervalle $[0,1]$

2.1.3. Les sous-ensembles Φ -flous.

La notion d'appartenance régissant la théorie des sous-ensembles Φ -flous présente toujours la caractéristique de gradation. Mais elle présente un autre avantage : celui de tenir compte de l'indétermination lors de l'évaluation du degré d'appartenance.

Ainsi, la fonction d'appartenance ne prend plus comme valeur un point de l'intervalle $[0,1]$ mais un intervalle compris dans $[0,1]$. On remarque que plus l'intervalle est grand, plus l'indétermination augmente. Ainsi l'intervalle $[0,1]$ signifie que l'indétermination est totale. De même, si on a l'intervalle $[a,a]$ avec a compris entre 0 et 1, c'est qu'il n'y a aucune indétermination.

On représente les sous-ensembles Φ -flous par un ensemble de couples où le premier élément est un composant du référentiel considéré et où le deuxième élément est un intervalle de $[0,1]$ donnant son degré d'appartenance par sa proximité de 0 ou de 1 et son indétermination par son étendue, par rapport au sous-ensemble Φ -flou considéré.

exemple : soit le référentiel E défini de la façon suivante

$$E = \{a, b, c, d\}$$

soit le sous-ensemble flou A défini de la façon suivante

$$A = \{(a, [0,0]); (b, [0.8, 0.9]); (c, [0.2, 0.2]); (d, [0,1])\}$$

La logique Φ -floue s'inspire de ces notions c.à.d. que ses valeurs de vérité sont des intervalles sur $[0,1]$

exemple : à la question "cette fille est-elle jolie ?", la logique Φ -floue permet, en plus de la gradation de notre avis sur $[0,1]$ de tenir compte de l'indétermination.

Ainsi, si à la question on répond

$[0.6, 0.9]$ fille assez ou même très jolie

$[0.4, 0.6]$ avis mitigé

$[0,1]$ indétermination totale (on ne connaît peut-être pas la fille en question)

Les variables ϕ -floues peuvent prendre les valeurs que peut prendre la fonction d'appartenance ϕ -floue c.à.d. tous les intervalles possibles dans $[0,1]$

Nous utilisons aussi des fonctions de variables ϕ -floues. Elles sont définies comme une application de F^n dans F (F étant l'ensemble des intervalles de la forme $[a_1, a_2]$ inclus dans $[0,1]$, n étant le nombre de variables ϕ -floues intervenant dans la fonction)

Les opérateurs reliant ces variables sont les opérateurs ϕ -flous MAX, MIN et complémentation.

exemple d'utilisation : $\text{Max} ([0.6, 0.9], [0.1, 0.7]) = [0.6, 0.9]$

le complément de $[0.3, 0.6]$ est $[1-0.6, 1-0.3] = [0.4, 0.7]$

Pour tout complément au sujet des sous-ensembles flous, je renvoie le lecteur au livre de Kaufman : (1973)

"INTRODUCTION A LA THEORIE DES SOUS-ENSEMBLES FLOUS -
TOME I ELEMENTS THEORIQUES DE BASE"

(Masson)

2.2. Mise en oeuvre des fonctions Φ -floues

La construction des fonctions Φ -floues à partir des fonctions booléennes citées antérieurement comporte deux étapes :

- l'affectation d'une valeur Φ -floue aux signes, symptômes et à l'index clinique
- l'extension des opérateurs booléens ET, OU et NON en opérateurs Φ -flous.

2.2.1. a. Affectation des variables Φ -floues

C'est à partir des réponses aux questions figurant dans le dossier thyroïdien actuel que nous allons affecter une valeur aux variables Φ -floues représentant l'interprétation des différents signes relevés.

Les variables que nous devons transformer correspondent aux questions de l'interrogatoire sur la Pathologie thyroïdienne (voir exemple de questionnaire en annexe).

Ces variables à transformer sont différentes selon leur type.

Ces différents types de variables booléennes sont les suivants :

- Questions de type OUI/NON/X sont représentées par 2 valeurs booléennes. La première prend pour valeur 1 si la réponse est OUI, 0 pour les autres réponses; la deuxième dite "variable d'existence" prend pour valeur 1 si la réponse est OUI ou NON, 0 si la réponse est "je ne sais pas" (X).
- Pour les questions de type "intensité", une variable booléenne est créée et prend pour valeur
 - 0 si l'intensité est 0 ou 1
 - 1 si l'intensité est 2 ou 3

Ici, aussi, une deuxième variable d'existence va indiquer si le signe a été ou non recherché.

- Pour ce qui est des questions de type "valeur", le signe est découpé grâce à l'introduction de bornes. On définit ainsi le plus souvent trois variables booléennes qui indiquent la situation de la valeur considérée par rapport aux bornes ;
 - Diminution : $X < B_1$
 - Valeur normale : $B_1 < X < B_2$
 - Augmentation : $X > B_2$
- De plus, une variable d'existence vient s'ajouter à ces vrais variables.

Ainsi par exemple, la valeur du taux de fixation à la 6e heure d'iode par la thyroïde, va correspondre aux trois variables booléennes mutuellement exclusives :

TF6 diminué : vrai si $TF6 < 15 \%$

TF6 normal : vrai si $15 \% < TF6 < 25 \%$

TF6 augmenté: vrai si $TF6 > 25 \%$

Une démarche analogue va nous permettre de déterminer la valeur des variables Φ -floues, qui sera :

- pour les questions OUI/NON/X

[0,1] si le signe n'est pas relevé ou inconnu (X)

[0,0] si la réponse est non

[1,1] si la réponse est oui

- pour les questions de type intensité, on convient d'attribuer à la variable Φ -floue correspondante, les valeurs suivantes :

[0,1] si le code est X

[0,0] " 0

[0,1/3] " 1

[1/3,2/3] " 2

[2/3,1] " 3

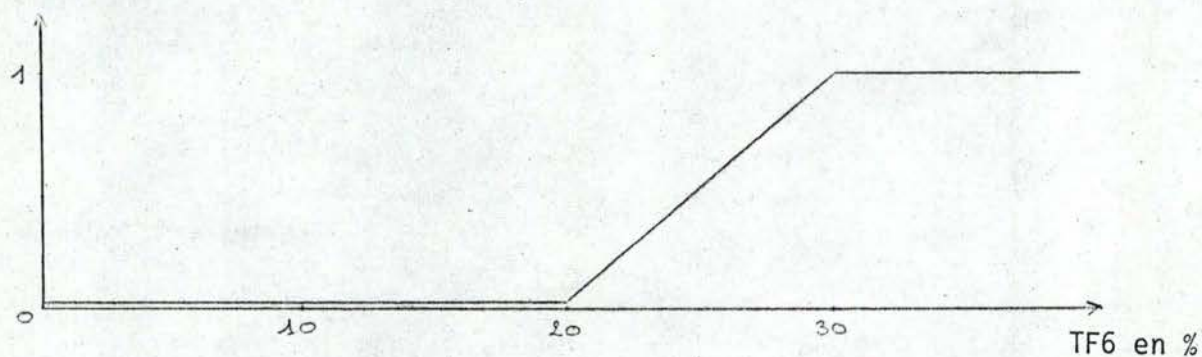
- En ce qui concerne les questions de type valeurs, nous pouvons, en prenant pour repères les bornes précédentes, définir les variables Φ -floues "valeur augmentée", "valeur normale", et "valeur diminuée".

Mais comme le recueil de données par le questionnaire ne nous donne pas d'indication sur la précision escomptée des différents dosages, il est donc impossible d'évaluer l'indétermination attachée aux variables Φ -floues correspondantes.

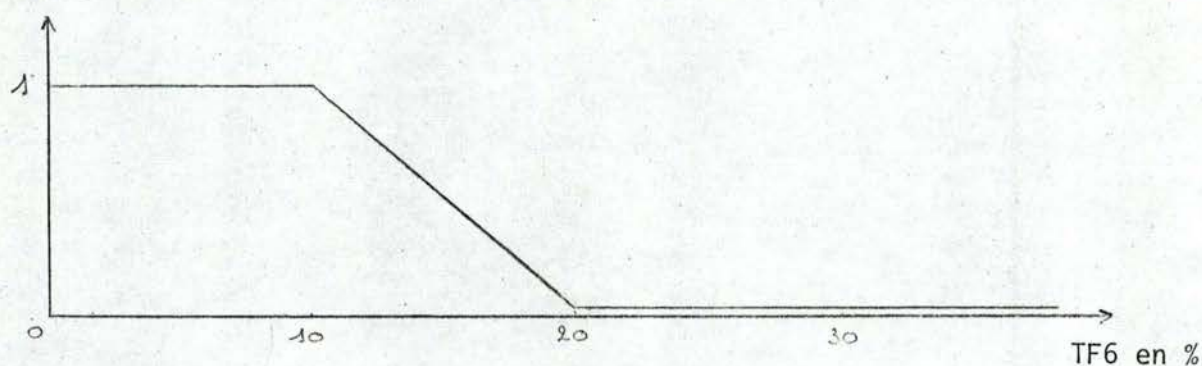
En conséquence, les valeurs de ces variables seront des intervalles réduits à des points, sauf dans le cas où l'examen n'est pas pratiqué, les trois variables étant alors égales à [0,1] .

En reprenant l'exemple du taux de fixation à la 6e heure, nous donnons ci-dessous l'illustration de trois ensembles Φ -flous TF6 augmenté, TF6 diminué et TF6 normal qui peuvent être construits sur la base des valeurs du taux de fixation. Nous avons choisi une simple représentation linéaire.

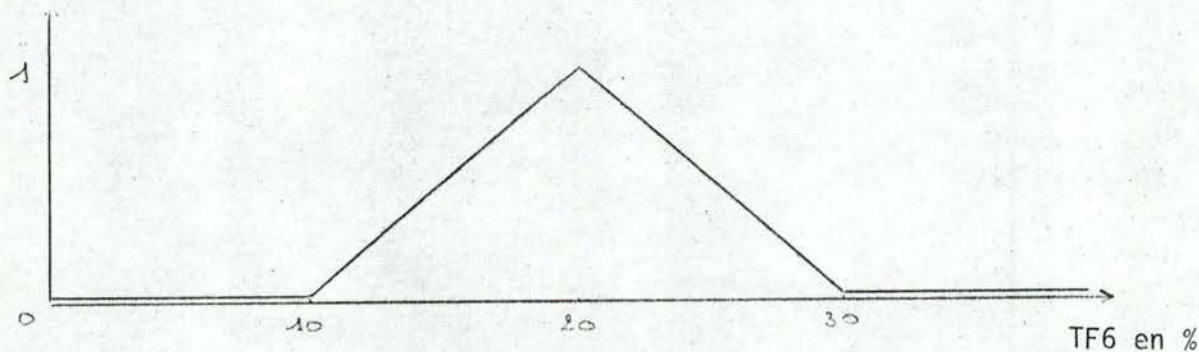
TF6 augmenté



TF6 diminué



TF6 normal



2.2.1. b. Transformation de l'index clinique

La construction adoptée pour donner une valeur Φ -floue à l'index clinique fait intervenir un opérateur Φ -flou de combinaison convexe des variables \tilde{x} -floues correspondant aux huit signes cliniques de l'index (voir index clinique en logique booléenne) :

Soit $[a_{1i}, a_{2i}]$ la valeur de la variable Φ -floue correspondant au signe i , l'Index prendra pour valeur l'intervalle :

$$I = \left[\sum_{i=1}^n w_i a_{1i}, \sum_{i=1}^n w_i a_{2i} \right] \text{ avec } \sum_{i=1}^n w_i = 1$$

Soit ici : $w_1 = w_2 = \dots w_n = 1/n$ avec $n = 8$.

Cet index Φ -flou conserve donc, de par sa construction, le caractère de simple sommation des pondérations des 8 signes, qui était propre à l'index clinique initial. Remarquons toutefois que son résultat va tenir compte, pour chaque malade, de l'intensité des différents signes, ainsi que des signes éventuellement non relevés, ces derniers contribuant alors à accroître l'indétermination (la longueur de l'intervalle) du résultat.

2.2.2 L'extension des opérateurs booléens.

Les opérateurs booléens à transformer sont le ET, le OU et la négation.

Les opérateurs \tilde{x} -flous choisis pour les remplacer sont :

Max. remplaçant le OU

Min. " ET

Non pour l'opération de complémentarité.

2.3. Quelques notions sur l'interprétation du résultat des fonctions Φ -floues.

Le résultat des fonctions Φ -floues se présente comme un intervalle Φ -flou (intervalle dans $[0,1]$). L'interprétation de cet intervalle se fait en terme :

- d'indétermination : fonction de la longueur de l'intervalle
- d'intensité : fonction de la situation de l'intervalle par rapport à 0 ou à 1.

Par exemple, les résultats suivants pour la fonction HYPER peuvent être traduits comme suit :

HYPER = [0,1] : impossibilité de préciser avec les signes dont on dispose, si le malade est hyperthyroïdien ou non.

HYPER = [0.7, 0.8] : hyperthyroïde franche car le niveau d'indétermination est faible et l'intervalle est proche de 1.

HYPER = [0.5, 1] : malade hyperthyroïdien, le degré de son hyperthyroïde ne peut pas être précisé avec les signes dont on dispose (degré d'indétermination est trop grand).

HYPER = [0.35, 0.45] : malade présenté comme normal mais présentant tout de même certains signes d'hyperthyroïde.

2.4. Résultats obtenus lors du passage en machine de ces fonctions Φ -floues.

Ces fonctions Φ -floues ne sont pas appliquées actuellement au C.H.U. Timone de Marseille où elles ont été élaborées et où le calcul flou est pourtant toujours à l'honneur.

Néanmoins, elles ont fait l'objet d'un important essai portant sur le dossier médical SYCVAR (1 700 dossiers thyroïdiens). Cet essai se limitait aux deux fonctions Φ -floues HYPER et HYP0.

Sur cette base, une étude comparative a été faite pour chaque dossier entre les résultats des deux fonctions Φ -floues et d'une part le diagnostic "machine" actuel comprenant le diagnostic d'orientation par les fonctions booléennes et le diagnostic d'affinement et d'autre part le diagnostic porté dans le service hospitalier.

Le diagnostic d'orientation Φ -flou n'est jamais en contradiction avec le diagnostic booléen. Les dossiers orientés hypo ou hyper par les fonctions booléennes présentent une valeur Φ -floue supérieure à [0.5,0.5]
Seuls 8 dossiers sur ceux diagnostiqués HYper présentent une fonction Φ -floue

HYPER un peu inférieure (entre 0.4 et 0.5). Cela est sans doute dû à l'index clinique ϕ -flou tel que nous l'avons défini. En effet, il ne constitue pas une stricte extension de l'index clinique antérieur et se trouve ainsi dans certains cas sous-évalué.

La possibilité de graduer l'importance des fonctions hyper et hypo dans $[0,1]$ permet, dès l'étape d'orientation, de distinguer les formes frustes des formes franches dans l'élaboration d'un diagnostic. Inversement, parmi les dossiers qui n'auraient pas été retenus par les fonctions booléennes on peut reconnaître ceux qui, bien que pouvant être considérés comme normaux, présentent une tendance hyper ou hypothyroïdienne (par exemple si hyper = $[0.3, 0.4]$).

Pour la plupart de ces dossiers, soit le diagnostic du service hospitalier, soit le diagnostic d'affinement mentionnent la présence d'une forme dite pré-pathologique.

Il semble donc que les gradations introduites par les fonctions ϕ -floues correspondent aux différentes nuances des états pathologiques considérés.

Des améliorations sont à envisager notamment l'adaptation du questionnaire d'entrée, en vue de pouvoir mieux recueillir la part d'indétermination qui peut s'attacher aux différentes données, tant cliniques que biologiques, ou encore l'introduction de nouveaux opérateurs autres que MIN et MAX.

Ces améliorations poursuivent toujours le même but : assister et approcher au mieux les raisonnements diagnostiques du médecin.

3. Analyse du type multidimensionnelle

Nous rappelons du point de vue théorique qu'étant donné des groupes prédéfinis D_i de points dans un espace à n dimensions, chaque point M étant déterminé par n composantes, il s'agit connaissant les n composantes d'un point quelconque, de déterminer le groupe auquel ce point appartient.

Pour notre domaine médical, les groupes D_i représentent les différents groupes diagnostiques, les points $M (X_1 \dots X_n)$ représentent les patients dont les signes caractéristiques sont les composantes du point.

Hypothèses de départ.

- Les points $M (x_1 \dots x_n)$ appartiennent obligatoirement à un seul des groupes
- M est le volume élémentaire d'espace entourant le point M
- $f_i(M)$ est la fonction de densité de probabilité connue de chaque groupe D_i
- on connaît la probabilité que M appartienne à dM si M appartient à D_i

$$P(M \in dM / M \in D_i) = f_i(M) dM$$

- on connaît la probabilité du point M d'appartenir à chacun des groupes alors qu'on ignore ses coordonnées; $P(M \in D_i)$ est une constante pour chaque groupe : soit $P(M \in D_i) = q_i$
- le médecin doit pouvoir attribuer à tout mauvais classement un coût.

L'analyse que nous opérons ici travaille en deux temps.

- 1° calculer la probabilité d'appartenance d'un point donné M à chacun des groupes D_i .
- 2° au vu de ces résultats, on décide à quel groupe D_i faut-il rattacher ce point M avec un minimum d'erreur possible.

1° Recherche des probabilités d'appartenance.

Nous devons rechercher la probabilité d'un point M dont nous connaissons les composantes à tous les groupes D_i ou encore :

$$P_i = P(M \in D_i / M \in dM)$$

Pour cela, on utilise le théorème de BAYES :

$$P(A/B) = \frac{P(B/A) \cdot P(A)}{P(B)}$$

or $P(B)$ peut se mettre sous la forme $P(B/A)P(A) + P(B/-A)P(-A)$

$$\text{d'où} \quad P(A/B) = \frac{P(B/A) \cdot P(A)}{P(B/A) \cdot P(A) + P(B/-A)P(-A)}$$

ou encore en prenant les A_i tels que l'intersection des A_i est vide et l'union des A_i est A (Partition de A), le dénominateur devient :

$$\sum_i P(B/A_i) P(A_i)$$

Ce résultat appliqué au calcul du P_i donne le résultat suivant

$$P_i = \frac{f_i(M) dM \cdot P(M \in D_i)}{\sum_j f_j(M) dM \cdot P(M \in D_j)} \quad \text{avec } D_i \cap D_j = \emptyset$$

ou $P(M \in D_i)$ est la probabilité à priori qu'un point M quelconque de l'espace appartienne au groupe D_i .

Tous les facteurs constitutifs de P_i sont connus, on peut donc calculer P_i .

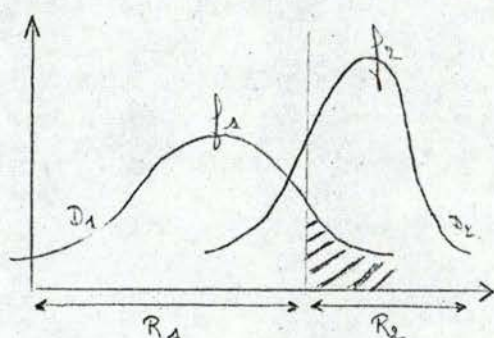
2° Attribution d'un point M à un groupe D_i .

Pour ce qui est du second problème, il faut décider en fonction des P_i obtenus, à quel groupe appartient ce point. C'est ici que s'introduit la notion d'erreur de classement.

Soient R_i les sous-ensembles obtenus par la partition de l'espace d'origine en groupes selon des surfaces à déterminer. Nous considérons qu'un point M appartient à la population D_i si et seulement si il appartient au sous-ensemble R_i . Dès lors, on peut commettre les erreurs suivantes:

attribuer à D_i un point de R_i appartenant à une autre population D_j ou ne pas attribuer à D_i un point extérieur à R_i qui lui appartient.

Prenons un exemple où existerait deux groupes et où on aurait deux



distributions unidimensionnelles à fonctions de densité gaussienne.

exemple : la probabilité que M appartenant à D_1 soit un point de R_2 se représente par la partie hachurée du graphe.

On peut en déduire la probabilité de mal classer un point M :

$$\int_{\bar{R}_i} f_i(M) dM = \sum_{k \neq i} \int_{R_k} f_i(M) dM$$

où \bar{R}_i représente le complémentaire de R_i

De plus, il faut tenir compte de la provenance à priori de D_i , soit q_i d'où la probabilité d'observer M appartenant à D_i et de mal le classer est

$$q_i \cdot \int_{\bar{R}_i} f_i(M) dM$$

et la probabilité des mauvais classements est

$$\varepsilon = \sum_{i=1}^k q_i \int_{\bar{R}_i} f_i(M) dM \quad (1)$$

On introduit également une notion de COÛT associée à chaque type de mauvais classement.

Exemple concret :

Supposons une surveillance de malade où divers signes sont régulièrement pris en considération - pouls (x_1)

- tension artérielle max (x_2)
min (x_3)

- température (x_4)

- rythme respiratoire (x_5)

Les points M que nous classerons sont donc composés de ces cinq paramètres (x_1, \dots, x_5)

Supposons également deux groupes diagnostiques - Etat satisfaisant
- Etat alarmant

Si à un moment donné, on doit classer un point M obtenu dans un de ces groupes, les conséquences sont totalement différentes suivant le type d'erreur commise. En effet, il est nettement moins grave de classer un point M d'un patient dans un état normal dans le groupe "Etat alarmant" que de placer un patient dans un état critique dans le groupe "Etat satisfaisant". On voit donc que les risques d'erreur sont totalement différents. C'est pourquoi nous introduisons cette notion de coût.

notation : $C(i \rightarrow j)$ est le coût d'un mauvais classement : un point appartenant à D_i classé dans le groupe D_j

Le coût de l'erreur ϵ que nous avons défini en (1) est donc

$$C(\epsilon) = \sum_i C(j \rightarrow i) q_i \int_{R_i} f_i(M) dM$$

Il nous reste donc à déterminer les K_i c.à.d. trouver la partition optimum afin que $C(\epsilon)$ soit minimum.

- Supposons d'abord que les coûts sont toujours les mêmes

Alors la partition est déterminée par

$$\min_R [C(\epsilon)] = \min_R \sum_i q_i \int_{R_i} f_i(M) dM$$

Pour un point M , la probabilité d'erreur est minimale si on le classe dans le groupe i pour lequel, si $C(j \rightarrow i) = \text{constante } \forall i, j$

$$q_i \int_{R_i} f_i(M) dM > q_j \int_{R_j} f_j(M) dM \quad \forall j \neq i$$

En minimisant l'erreur sur chaque point, on la minimise sur l'ensemble. La partition s'effectue donc, comme on pouvait le supposer, par les surfaces déterminées par

$$q_i \int_{R_i} f_i(M) dM = q_j \int_{R_j} f_j(M) dM$$

- Supposons maintenant qu'on ait des coûts différents suivant les erreurs de classement.

L'introduction des coûts non égaux revient à déplacer la surface vers l'un des points moyens de chaque groupe.

Remarque ; on remarque que cette méthode classe obligatoirement tout point de l'espace, même s'il appartient aux zones de recouvrement diagnostique c.à.d. les zones limites entre deux ou plusieurs groupes diagnostiques.

Cette méthode a été appliquée au problème particulier de classement des observations de malades thyroïdiens dans un des quatres groupes diagnostiques habituels.

Le fichier de travail sur lequel la méthode a été testée comportait 654 dossiers.

Pour tester cette méthode, on avait besoin de choisir un ensemble de variables dont les valeurs particulières constituent un point M de l'espace correspondant à un patient.

A titre indicatif, neuf variables ont été choisies : pouls; tension artérielle maximale - minimale; rapport poids sur taille; taux de fixation du radio-iode aux 6e et 24e heures; fractions I, II et III du test Blanquet-Meyniel.

Pour pouvoir évaluer les résultats de la classification faite par une telle méthode, on a comparé avec les diagnostics établis par le médecin consultant lui-même.

Les résultats obtenus n'ont guère été satisfaisants : 65 % seulement des observations étaient bien classées. Ceci est inhérent à la remarque faite plus haut.

L'équipe de recherche n'a pas pour autant rejeté cette méthode. En effet, la classification est intimement liée au choix des caractères associés au patient.

C'est donc en recherchant des coefficients discriminants caractérisant mieux chacune des populations qu'il a été permis d'améliorer les résultats pour finalement aboutir aux chiffres suivants.

Groupes	Effectifs	nombre de dossiers bien classés	nombre de dossiers mal classés	% de dossiers bien classés
Hypothyroïde	25	18	7	72
Fonctions thyroïdiennes normales	448	367	81	82
Hyperthyroïde	133	111	22	83.5
Thyréose	48	27	21	56.2

On remarque que le pourcentage de dossiers bien classés est assez élevé en moyenne. Mais en médecine, ils sont insuffisants. Dans le domaine médical, on ne joue pas avec des pièces de machine mais avec des hommes. (irréversabilité des conséquences)

Les résultats cités ci-dessus ne nous ont pas paru suffisants. De plus, lorsque l'on calcule les différentes probabilités d'appartenance, on constate que la quasi-certitude (c.à.d. $P_i > 0.8$) est rarement atteinte.

C'est ainsi que l'équipe de Marseille a été amenée à tester la robustesse de cette méthode de classification. La méthode a été testée en modifiant encore les caractères servant à situer l'individu, mais très peu d'améliorations ont été obtenues.

On comprend donc que cette méthode qui peut conduire à des erreurs de classement même si la probabilité d'appartenance à un groupe est forte, entre en contradiction avec l'esprit médical.

Néanmoins, le choix d'un type de méthode étant fortement lié au problème à résoudre, cette démarche ne peut être rejetée.

4. Classement par pondération des caractères qualitatifs

La démarche étudiée sous ce point 4 est une illustration des méthodes de discrimination non paramétrique.

Cette méthode utilise exclusivement des caractères qualitatifs sous la forme de variables logiques.

Nous verrons au cours de cette explication que le grand avantage de ce genre de méthode est la simplicité.

Explication de la méthode.

Chaque caractère qualitatif étudié peut offrir trois possibilités :

- le signe est positif
- le signe est négatif
- le signe est douteux ou non relevé.

On veut affecter le même poids aux signes qu'ils soient positifs ou négatifs. Un moyen pour le réaliser est de dédoubler chaque signe en un signe positif et un signe nié.

Exemple : Le signe "toux du malade" est dédoublé en "toux du malade" et "non toux du malade".

- Si un patient tousse, le signe "toux du malade" sera égal à 1 et le signe nié sera nul.
- Si un patient ne tousse pas, le signe "toux du malade" sera nul et le signe "non toux du malade" sera mis à 1
- S'il y a doute sur ce symptôme ou s'il n'a pas été relevé, les deux signes seront nuls.

Supposons : S_h : signe d'un patient où h : varie entre 1 et 2 n car les n signes possibles sont dédoublés.

D_i : groupe diagnostique i où i varie de 1 à d (d étant le nombre de diagnostics possibles)

P_h : fréquence relative du signe S_h quel que soit le diagnostic

$P_{h/i}$: fréquence relative du signe S_h à l'intérieur du groupe diagnostique D_i

Tout malade est alors représenté par une série de 0 et de 1

$$\text{malade : } M = (\delta_1, \delta_2, \dots, \delta_h, \dots, \delta_{2n})$$

où $\delta_h = 1$ si le malade vérifie le signe

$\delta_h = 0$ sinon.

La méthode est basée sur le rapport suivant :

$$\frac{P_h}{P_{h/i}} \quad (1)$$

Si par exemple le signe h n'est observé sur la totalité du fichier que pour le diagnostic i , alors la valeur du rapport (1) est très grande pour tous les autres diagnostics et offre la plus petite valeur pour le diagnostic i (proche de 1)

On voit aisément qu'il a suffi alors de créer une fonction F pour chaque diagnostic i de la forme :

$$F_i = \sum_{h=1}^{2n} \frac{P_h}{P_{h/i}} \delta_h \quad (i=1, d)$$

Rm. δ_h affecte un poids nul aux signes douteux ou non relevés.

On obtient ainsi d valeurs de F_i pour un malade, on retiendra alors le diagnostic pour lequel F_i est minimum.

Application Thyroïde - Résultats.

Les P_h et les $P_{h/i}$ ont été calculés sur base d'un fichier comportant 1200 dossiers. Ces P_h et $P_{h/i}$ sont calculés une fois pour toute.

On a choisi non plus les 4 groupes diagnostiques traditionnels mais on les a subdivisés : 10 groupes (d=10)

Le nombre de signes et symptômes relevé par le questionnaire est de 144 (n=144)

Voici les résultats :

Hyperthyroïde

Diagnostics	Effectifs	Bons classements	%
Psychonévrose avec Hyperthyroïde	29	23	79
Maladie de Basedow	115	90	78
Goître basedowifié	12	10	83
Adenome toxique	17	17	100
Cardiothyroïse	10	10	100
Total : Hyperthyroïde	183	150	82

Thyréose

Diagnostics	Effectifs	Bons classements	%
Goître avide d'iode	23	19	82

Fonctions thyroïdiennes normales

Diagnostics	Effectifs	Bons classements	%
Goître simple diffus	88	79	89
Goître simple modulaire	73	70	96
Psychonévrose avec enthyroïde	179	160	88
Total : normaux	340	309	91

Hypothyroïde

Diagnostics	Effectifs	Bons classements	%
Hypothyroïde	39	20	51

Les chiffres de ces différents tableaux montrent que la fraction de malades bien classés est élevée.

Toutefois cette méthode offre un inconvénient grave : un signe rarissime peut ne pas avoir été observé pour un diagnostic au sein de l'échantillonnage ce qui entraînera par la suite pour tout malade qui offrirait ce signe, un terme correspondant infini, ce malade ne pourra alors être classé comme provenant de ce diagnostic auquel il appartient pourtant. Pour la même raison, une erreur commise sur la collecte d'un seul signe peut avoir des conséquences semblables. C'est ce qui rend cette méthode fragile. De plus, comme pour la méthode d'analyse discriminante de Fisher, il existe des mauvais classements.

5. Conclusions sur ces différentes applications

On peut créer un tableau comparatif entre les différentes méthodes employées pour le diagnostic relatif à la Pathologie thyroïdienne.

Groupes	% de dossiers bien classés		
	booléen	discriminante	qualitatif
Hypothyroïde	89	72	82
Normal	91	82	82
Hyperthyroïde	91	83.5	91

On voit que les meilleurs résultats sont obtenus par la méthode booléenne.

La méthode la moins bien adaptée à la pathologie thyroïdienne est la méthode de type : analyse multidimensionnelle. Comme nous l'avons évoqué, le grand avantage de la méthode logique est la quasi-certitude du diagnostic répondant à l'une des fonctions. Si aucune des fonctions n'est satisfaite il existe une branche de rejet. Donc, contrairement à l'analyse discriminante, tout patient n'est pas nécessairement classé.

Mais ceci ne veut pas dire que la méthode des fonctions booléennes est la meilleure. Nous disons qu'elle est le mieux adaptée au problème de la pathologie thyroïdienne. La méthode discriminante sera mieux adaptée dans d'autres problèmes où interviennent de nombreuses variables continues.

Il n'existe donc pas en informatique médicale d'algorithme miracle qui pourrait être adapté à toute pathologie. En effet, toute pathologie possède ses caractéristiques propres qui dirigeront le choix de la méthodologie.

Nous ne nous posons pas ici la question de la portée et de l'avenir du diagnostic automatique. Nous en discuterons lors de la conclusion de ce travail.

DEUXIEME PARTIE

A - ÉTUDE DU DOSSIER MÉDICAL : SYCVAR

A. ETUDE DU DOSSIER MÉDICAL : SYCVAR

Le dossier médical élaboré par l'équipe du CHU Timone de Marseille se nomme SYCVAR.

C'est une étude de ce modèle que nous allons entamer.

°
° °

CHAPITRE 1

INTRODUCTION

La prise d'informations a toujours été un problème crucial en Informatique. Le problème est plus vif encore lorsqu'il se pose dans le domaine médical. En effet, il peut exister certains "trous" d'informations très nuisibles à l'exploitation des documents (ex. : impossibilité de comparer deux malades, critère par critère si certains signes sont absents pour l'un d'entre eux).

Un autre problème posé par le monde médical est le manque de standardisation du vocabulaire. Derrière un même mot peuvent se cacher plusieurs significations. Par ailleurs, même lorsque la sémantique d'un mot a été acceptée, le facteur d'appréciation d'un signe ou d'un symptôme est très personnel et donc très fluctuant.

Avec l'apparition des traitements électroniques des informations, l'idée d'automatiser, de systématiser le recueil des informations médicales a germé dans beaucoup d'esprits.

Pour remédier aux problèmes évoqués plus haut, l'équipe du CHU Timone a opté, lors de l'élaboration du DOSSIER MEDICAL, pour recourir à un questionnaire demandant des réponses strictes et précises (cfr. l'exemple en annexe 1).

Ce type de questionnaire répond en fait à plusieurs critères :

- qu'il était nécessaire d'imposer au personnel médical la prise d'observations fiables et complètes;
- qu'il était nécessaire de systématiser la sémantique médicale;
- qu'il était nécessaire que les observations soient les plus complètes possibles, mais soient sobres et faciles à recueillir;
- qu'il était nécessaire de pouvoir exploiter ces données facilement et de manière "scientifique".

L'élaboration même de ce questionnaire a un apport bénéfique sur la pathologie traitée étant donné qu'il demande au médecin une remise en question de ses connaissances, une vision claire et précise de la maladie étudiée. Ce type de questionnaire pose également certaines contraintes d'utilisation. Il s'agit surtout d'un problème d'attention et d'application lors du remplissage du questionnaire : respecter l'ordre et la sémantique.

L'utilité d'un tel système porte sur trois points :

- a. Un tel dossier garde la "trace" d'un malade après son passage dans un service (situation du malade, particularités de l'évolution, thérapeutique adoptée).

Le système offre ainsi divers services à l'utilisateur.

Exemple : - accéder à tout instant aux données du malade

- éditer un résumé

- éditer une lettre de synthèse au médecin traitant.

- b. Un tel dossier aide la recherche médicale. Ainsi, il peut aider à l'amélioration de la démarche diagnostique en mettant en évidence le "poids" diagnostique de certains signes, dosages, ...
- c. Un tel dossier aide également la gestion administrative des hôpitaux en recueillant tous les actes de tous les corps hospitaliers. En effet, les actes médicaux génèrent toujours des séries d'"actes-conséquences" administratifs.

°

° °

CHAPITRE 2

PRÉSENTATION DU SYSTÈME

Le but de SYCVAR est de proposer une forme de rédaction (questionnaire) qu'il est capable d'entrer en machine et d'exploiter.

2.1. QUELLES SONT LES INFORMATIONS RETENUES PAR SYCVAR ?

Lors de réponses à un questionnaire, plusieurs types de réponses sont possibles. Le système doit en tenir compte. SYCVAR accepte 5 types de données.

- Les questions du type "Existe-t-il un goître ?" appellent des réponses dites logiques, c'est-à-dire "oui" ou "non" ou encore "je ne sais pas".

- Ces mêmes questions peuvent proposer des réponses pondérées:

Exemple : pour le goître

<u>Intensité</u> du signe :	- un lobe palpable	0
	- isthme palpable	1
	- isthme et lobe palpables	2
	- gros goître	3

- Les nomenclatures sont fréquentes : nous prendrons alors un code diagnostique ou bien une codification d'origine ethnique

Exemple : demande

- de la profession du patient
- de traitements déjà reçus
- ...

Il s'agit en fait de valeurs numériques discrètes que nous appellerons "code".

- Les résultats de laboratoires sont très souvent des mesures, il s'agit de valeurs numériques le plus souvent continues.

Exemple : en irathérapie

- surface scintigraphique en cm^2
- poids de la thyroïde en grammes
- période effective en jours
- ...

- On peut aussi vouloir ajouter des remarques, du texte pour une retranscription ultérieure.

Exemple : après une question qui demande comme réponse OUI, NON, on demande des précisions.

- Le malade se plaint-il d'ATHERO-SCLEROSE CEREBRALE ?
(pertes de mémoire, affaiblissement psychique ...)
Précisez lesquels :

Un sixième type de variable peut apparaître dans le dossier : il s'agit des variables générées c'est-à-dire variables dont la valeur est automatiquement calculée à l'aide de valeurs déjà connues tels les rapports, poids sur tailles et les indices divers.

Remarque : il est très important de faire la différence entre une réponse "non" et une réponse "je ne sais pas".

2.2. STRUCTURE LOGIQUE

Le dossier médical du système SYCVAR se décompose en deux parties distinctes : d'une part le dossier de base (document constituant une image instantanée de l'état pathologique du malade à son entrée à l'hôpital); d'autre part, des documents de mise à jour reflétant l'évolution du malade ou des documents de complément clinique ou paraclinique, ce sont les dossiers satellites.

DOSSIER DE BASE

Organisation : - identification administrative et technique du malade
 - antécédents (hérédité)
 - interrogatoire sur les signes ayant pu apparaître
 - examen clinique
 - examens paracliniques : type de dosage biologique
 - diagnostics
 - (thérapeutique).

La construction d'un tel questionnaire est très importante et doit être faite en plusieurs étapes.

1ère étape

- établir une liste de symptômes et de signes (établie par le clinicien)
- fractionner la pathologie en chapitres autonomes formant une partition.

La pertinence de ce fractionnement jouera un rôle déterminant dans la qualité des documents.

2ème étape

Critique de cette liste afin de :

- éliminer les questions purement descriptives
- choisir le type de chacune des questions
- définir les questions commandes, c'est-à-dire les questions influant sur les réponses des questions suivantes.

Il en existe deux sortes :

1. Commandes parenthésées

Il s'agit de différents chapitres imbriqués pour lesquels la première question du chapitre implique dans certains cas, une réponse systématique aux autres questions.

2. Commandes à distance

Il s'agit de variables dont la réponse conditionne la réponse à une ou plusieurs variables situées plus loin dans le dossier.

Exemple : la question "sexe du patient ?" influera sur le choix de têtes de chapitre : examen gynécologique homme ou examen gynécologique femme.

- Mettre au point la typographie du questionnaire :
 - le rendre facile à remplir
 - le rendre facile à relire
 - le rendre facile à perforer ou à taper sur visu.

3ème étape

Tester le questionnaire en testant par exemple sur une cinquantaine de cas.

4ème étape

Le recueil en routine.

DOSSIERS SATELLITES

Ce sont les dossiers de complément du type clinique et/ou de type paraclinique. Ils sont nécessaires pour compléter les dossiers de base d'une part trop importants et d'autre part pouvant se répéter sans avoir recours de nouveau à un dossier complet.

Il existe aussi des dossiers satellites de mise à jour : ces dossiers sont appelés "fiches-suites" et sont pris au fur et à mesure du déroulement de la maladie pour en refléter fidèlement l'évolution.

2.3. L'ENTREE DES INFORMATIONS

Les informations sont lues et classées par type : logique, intensité, code, valeurs numériques et textes.

Les contrôles effectués sont les suivants :

- contrôles formels : des caractères et de l'enchaînement des réponses, en particulier les questions dont les réponses dépendent des réponses antérieures.

- contrôles des codes : chaque code inscrit doit figurer dans la liste de référence.
- contrôles de vraisemblance logique dans certaines associations de signes proposées par les cliniciens ou sélectionnées par les premières exploitations.

Lorsque les informations sont ainsi validées par la machine, elles sont compactées (gain de place) et enregistrées sur disque magnétique (accès direct).

Une table de chaînage est mise à jour, elle permet de relier le dossier de base aux différents dossiers satellites recueillis pour un sujet donné.

2.4. LES EXPLOITATIONS

On peut distinguer deux catégories principales.

1. Les exploitations traitant les sujets un à un.

Par exemple : - édition du résumé de dossier
 - édition de la lettre-réponse au médecin traitant.

2. Les exploitations traitant des ensembles de sujets, qui sont les plus nombreuses.

Par exemple : - les dénombrements
 - les tris
 - les moyennes écart-types, test de X^2 , ...

Ces exploitations simples sont disponibles en temps réel car elles n'exigent pas de méthodologie statistique trop évoluée.

Lorsque les études de type 2 vues précédemment mettent en jeu une méthodologiestatistique plus élaborée, l'exploitation nécessite un interface entre les fichiers contenant les dossiers et des fichiers de travail. Ces derniers ne contiendront que les informations à traiter, classées et aucune autre.

Cette tâche est assumée par des programmes de préparation existant au CHU Timone de Marseille sous le nom de PREPAR ou PRESTAT.

Sur ces fichiers de travail, on pourra obtenir des exploitations plus élaborées telles que :

- Editions de fréquences, de moyennes et écart-types pour des populations de sujets répondant à des critères définis par les cliniciens.
- Détection à intervalles réguliers des variations significatives de certains dosages, signes, associations de signes pour des groupes pathologiques donnés.

- Ventilation des entrants par diagnostic (ou par critère plus élaboré)
- Taux de rotation des malades

Certaines exploitations plus particulières peuvent encore être obtenues à la demande.

- Analyse multivariée :

- descriptive : type principaux composants, ...
- explicative : type régressions multiples, analyses factorielles
- classification : type analyse discriminante de Fisher ...

- Analyse logique :

- explicative : recherche des Minterms impliquant des diagnostics
(recherche des associations de signes pathognomiques)
: axiomatisation de la connaissance médicale
- classification : aide au diagnostic logique.

o

o o

CHAPITRE 3

SPÉCIFICATIONS PLUS TECHNIQUES A PROPOS DU SYSTÈME SYCVAR

3.1. GENERALITES

La "variable" c'est-à-dire la réponse à une question constitue l'information élémentaire.

Néanmoins, l'unité logique de traitement est l'unité dossier d'un sujet, c'est-à-dire tous les éléments qui permettent de gérer et d'exploiter les informations contenues par le questionnaire médical (dossier de base + satellites).

Nous avons déjà vu qu'il existait un dossier de base et des dossiers satellites. Le lien entre ces éléments est constitué par un fichier "CATALOGUE" qui identifie le dossier de base et chacun des dossiers satellites grâce à leurs paramètres spécifiques.

Les questionnaires entrent en machine et sont mémorisés sous forme condensées sur des fichiers de stockage, un par nature de dossier. Ces fichiers ne sont accessibles que par le programme d'entrée, de mise à jour et par les programmes d'exploitation.

Remarque : il s'agit de fichiers de référence nécessitant des protections contre les fausses manoeuvres et les accès non autorisés.

3.1.1. Les informations élémentaires

La structure du questionnaire permet d'organiser les données séquentiellement, c'est-à-dire qu'un item est repéré tant sur le questionnaire que dans l'enregistrement magnétique par son ordre d'apparition, c'est-à-dire par un numéro.

La sémantique d'un terme comporte deux éléments : celui qui identifie l'item, celui qui supporte sa valeur.

- L'identification se fait par le numéro d'ordre qui est implicite
- La valeur de l'item est donc seule rentrée en machine.

Un enregistrement correspondant à un dossier est donc une suite ordonnée de valeurs.

3.1.2. Notions d'existence de l'information élémentaire

Chaque variable est dédoublée à son entrée.

La première variable résultant de ce dédoublement correspond au type de la question (logique 0,1 ; intensité 0,1,2,3; ...).

La deuxième est générée automatiquement et vaut :

- 1 si l'information existe
- 0 sinon.

Ce dédoublement est utile pour ne pas confondre une variable nulle (0,1) avec une variable "je ne sais pas" (0,0).

3.1.3. Hiérarchie des informations

Les unités-dossiers (dossier de base + dossiers satellites) sont immatriculées dans un fichier "CATALOGUE" qui comporte les coordonnées des fichiers supportant la description de chaque dossier élémentaire constitutif de l'unité-dossier considérée. Ces fichiers de description sont appelés "REPERTOIRE". A chaque répertoire correspond un dossier élémentaire.

Le "CATALOGUE" pointe donc vers les différents "REPERTOIRES" dont l'ensemble décrit les caractéristiques de l'unité-dossier en cause.

A son tour, le "REPERTOIRE" pointe vers le fichier de stockage du dossier considéré (données médicales) et vers les différents fichiers annexes qui décrivent ce dossier et permettent sa manipulation (paramètres, tables de correspondance, ...). (Cfr. Schéma page 83).

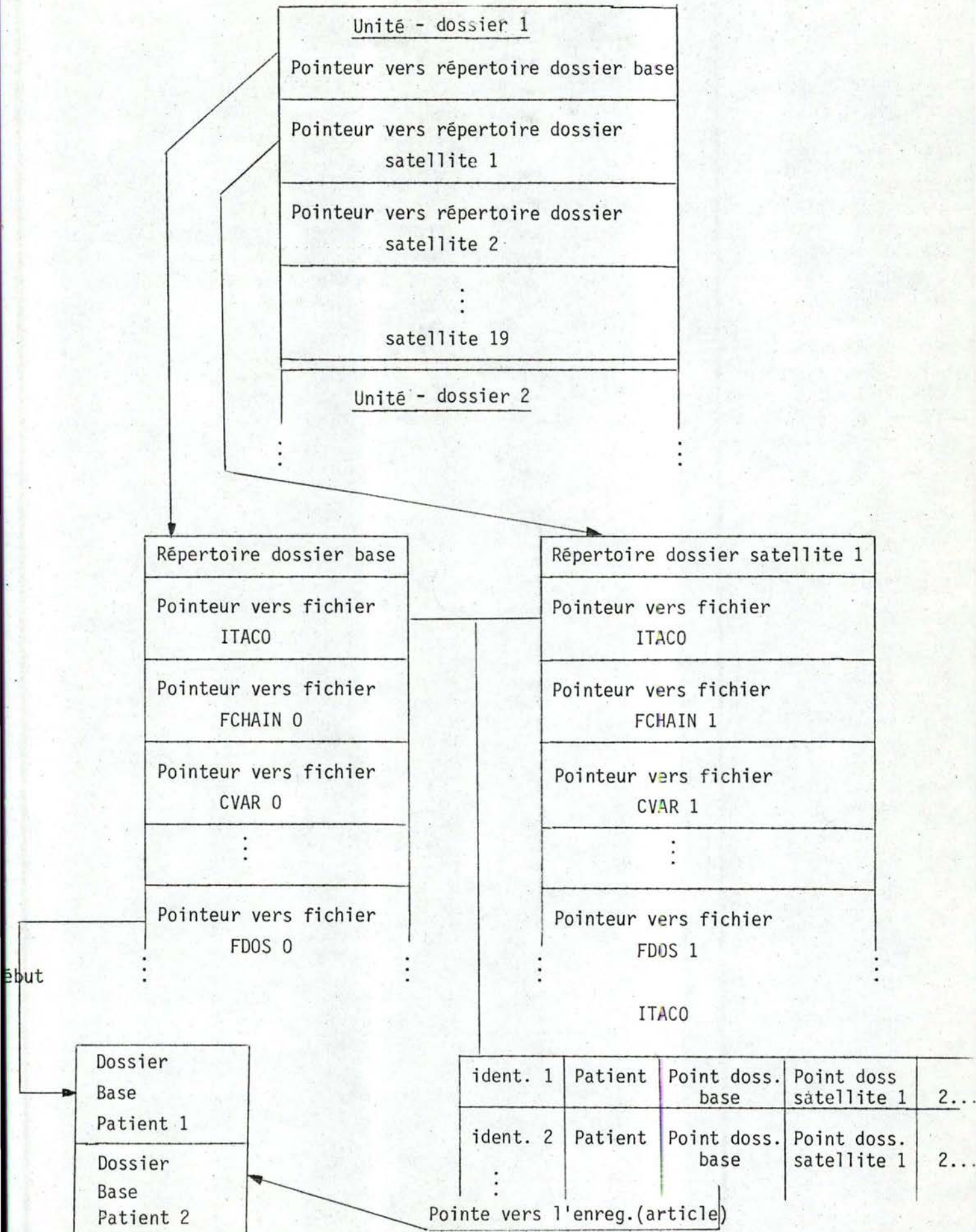
3.1.4. Fichiers descripteurs pointés par l'intermédiaire du répertoire

Le "REPERTOIRE" comprend les noms et les caractéristiques des fichiers descripteurs.

Ces fichiers sont cités ci-dessous avec leur utilité.

- ITACO : fichier commun à tous les répertoires puisqu'il n'en existe qu'un seul pour l'ensemble unité-dossier (base + satellite). Il s'agit d'une table qui identifie chaque patient et comporte les pointeurs permettant de retrouver tous les dossiers satellites (dossier de base compris).
- FCHAIN : fichier permettant de chaîner les satellites entre eux.
- CVAR : fichier décrivant la structure de l'information médicale.
- CV : fichier dans lequel se trouve la liste des moments autorisés pour les différents codes de l'observation. Pour chaque valeur du code, on trouve le texte en clair correspondant.
- IMAGE : fichier permettant, pour le dossier en cours de remplissage, de connaître l'état exact des réponses déjà faites, demandées et non reçues, non demandées.
- FDOS : fichier de stockage des observations médicales sous forme condensée.

CATALOGUE



- TEXTR : fichier des textes en clair correspondant à chaque item du dossier et prévu pour l'édition des résumés pour les médecins traitants.
- FCOR : fichier qui comporte l'identification des fonctions de cohérence ainsi que des variables intervenant dans ces fonctions.
- TABACS : fichier comportant les différentes tables d'accès possibles pour retrouver les patients individuellement.
- PRETRT : fichier de description des fonctions de prétraitement et des variables qui les composent.
- GRILLE : fichier des différentes formes d'édition des résumés d'observation.
- CLEF : fichier comportant la liste des clefs d'accès individuelles et des niveaux d'accès autorisés pour chaque personne physique.

3.2. NOTIONS D'IMPLEMENTATION DE CES FICHIERS

ITACO

- Nombre d'enregistrements : il est égal au nombre maximum de patients prévus pour une unité-dossier.
- Longueur de l'enregistrement : elle est fonction du nombre d'octets nécessaires pour l'identificateur.
- Composition de l'enregistrement : en plus de l'identificateur permettant de caractériser le patient, existe le nom du patient suivi de 20 pointeurs :
 - un vers le dossier de base
 - les autres vers les dossiers satellites.

Le numéro de ligne de cette table est égal au numéro du dossier séquentiel du patient augmenté de 1 car le premier enregistrement est constitué d'un en-tête.

FCHAIN

- Nombre d'enregistrements : il est paramétré à la création en fonction du nombre de lignes de la table ITACO.
- Longueur de l'enregistrement : elle est dépendante de la fonction de chaînage.
- Composition de l'enregistrement : un enregistrement comporte le numéro de la ligne correspondante d'ITACO, suivi du pointeur vers l'enregistrement correspondant de l'observation médicale dans le fichier FDOS, enfin la fonction de chaînage.

Il s'agit d'une expression sous forme de chaîne d'octets, permettant de repérer et d'identifier les différentes occurrences d'un même dossier satellite pour un même patient.

Remarque : il existe un nombre maximum d'occurrences d'un dossier satellite donné pour un patient donné.

CVAR

- Nombre d'enregistrements : il est égal au nombre maximum de variables du dossier + 1. Car le premier enregistrement contient des données générales décrivant, à la fois, le fichier et la forme globale de l'observation.
- Composition de l'enregistrement différent du premier.
 - type de la variable.
 - son numéro dans le type.
 - les extrêmes autorisés de la valeur s'il s'agit d'une variable numérique, ou l'indice de début de la variable texte si le type est égal à 5.
 - un drapeau indiquant si cette variable est une commande.
 - si cette variable est commandée, on trouve ensuite le n° de la variable (n° de ligne de table CVAR) qui commande à distance la variable en cours.
 - le codage possible en fonction de la commande à distance.
 - le n° de la variable (de la variable en cours dans le cas de commande de type parenthésée.
 - le codage possible pour la valeur de la commande de type parenthèse sur les 2 mots suivants, on trouve les niveaux d'accès autorisés en entrées et en sorties.
 - la longueur en nombre de caractères de la variable en cours de traitement.
 -
 - pointeurs vers fichiers de cohérence, fichier de pré-traitement.
 - texte interrogatif qui apparaîtra sur la visu lors du programme d'entrée des données.

CV

- Nombre d'enregistrements : il est fonction du nombre de codes et de moments.
- Composition de l'enregistrement : le premier est différent des autres et est constitué d'un en-tête.

Les autres comportent des pointeurs permettant de retrouver tous les moments et tous les textes relatifs à un code.

Composition

- n° d'ordre du code.
- un indicateur égal à zéro si ce code est numérique; égal à 1 s'il est alphabétique.
- adresse de début des enregistrements comportant les moments possibles pour ce code.
- adresse du dernier enregistrement relatif à ce code.
- nombre de moments.
- adresse du premier des enregistrements contenant les textes en clair de chacun des moments.

IMAGE

- Nombre d'enregistrements : il est fonction du nombre d'observations médicales considérées comme "actives". On appelle dossiers "actifs" des dossiers qui sont en cours de remplissage. Leur nombre dépend donc du débit du service clinique considéré.
- Longueur de l'enregistrement : elle est fonction du nombre de variables du dossier.
- Composition de l'enregistrement : le dossier IMAGE est condensé. Les valeurs de chaque élément correspondant à l'état de remplissage d'une variable sont comprise en 0 et 3 (codé sur 2 bits).
 - 00 : variable a été complètement remplie
 - 01 : question a été posée et la réponse est en attente
 - 10 : variable n'a pas été remplie.

C'est grâce à ce fichier que le programme d'entrée permet le redémarrage de l'entrée d'un dossier c'est-à-dire permet de compléter un dossier dont on a interrompu la saisie.

FDOS

C'est le fichier de stockage sous forme condensée des observations médicales.

- Nombre d'enregistrements :
 - Pour le dossier de base, le nombre d'enregistrements est égal au nombre de lignes de la table ITACO c'est-à-dire au nombre maximum de patients prévus pour cette unité-dossier.
 - Pour les dossiers satellites, il est estimé en fonction de la fréquence d'existence du satellite de nature donnée, ainsi que du nombre d'occurrences possibles pour un patient donné de ce satellite.

- Longueur de l'enregistrement : elle est dépendante du nombre de variables du satellite considéré.
- Composition de l'enregistrement : chaque enregistrement commence par trois mots réservés :
 - le premier : pointeur vers la ligne correspondante de ITACO
 - le second : pointeur vers le fichier IMAGE si le fichier est actif, sinon ce mot est à zéro.
 - le troisième : drapeau comportant la mention valide ou pas.

TEXTR

- Nombre d'enregistrements : autant que de variables plus un réservé pour l'en-tête.
- Longueur de l'enregistrement : déterminée lors de la définition des fichiers par les cliniciens.
- Composition de l'enregistrement : fichier composé des textes affirmatifs correspondant à chaque variable du dossier.

FCOR

Utilité

Ce fichier comporte la description des fonctions de cohérence c'est-à-dire des fonctions dont le calcul est déclenché automatiquement lorsque la dernière variable intervenant dans cette fonction est remplie. Cette fonction a pour but de vérifier la cohérence des données et de signaler toute incohérence au moment du remplissage.

- Composition de l'enregistrement : il est composé en deux parties :
 1. description des fonctions (résultats de l'analyse d'une formule)
 2. repérage des variables intervenant dans chacune des fonctions.

TABACS

Ce fichier est de format variable puisqu'il contient les tables de correspondance entre des clés de tri définies par le service clinique et les numéros de lignes correspondant à chaque patient dans la table ITACO.

Les plus fréquemment retrouvées sont des tables d'accès comportant comme clés de tri le nom et le prénom, les tables d'accès comportant le numéro d'identification administrative, les tables d'accès par code diagnostique.

PRETRT

Il a une structure identique au fichier FCOR. Les fonctions de pré-traitement ne donnent lieu à des calculs qu'au moment de l'enregistrement du dossier. Les calculs sont déclenchés au moment où le dossier est enregistré sur disque.

Ces fonctions sont destinées, en particulier, à tenir à jour les tables d'exploitation, telle la table de fréquence d'apparition de tel ou tel signe, ou des tables de moyennes de telle ou telle variable numérique. On peut constituer ainsi des fichiers de travail dont les renseignements sont mis à jour au fur et à mesure de l'entrée de nouveaux dossiers.

GRILLE

- nombre d'enregistrements : 6 enregistrements + 1 en-tête.

- Composition des enregistrements utiles :

Chaque enregistrement correspond à une grille possible de résumés pour le dossier considéré.

Chaque enregistrement correspond donc à une suite d'éléments liés à chacune des variables du dossier.

- Cet élément est égal à 1 si l'on veut que la variable correspondante apparaisse sans condition sur le résumé;
- cet élément est égal à 2 si l'on veut que la variable correspondante apparaisse soit, lorsque le signe est positif, s'il s'agit d'une variable logique, soit si elle a été recueillie, s'il s'agit d'une variable numérique;
- cet élément est égal à 3 si l'on veut que la variable correspondante soit ignorée.

CLEF

- Ce fichier ne comporte qu'un seul enregistrement.

- Chaque clé se compose de 8 caractères alpha-numériques définissant la clé plus 2 mots entiers correspondant au niveau d'accès (niveau d'accès en Entrée et en Sortie).

- Trois niveaux de clés :

- au niveau de l'unité-dossier;
- au niveau de chaque information élémentaire (niveau d'accès en entrée de la donnée et en consultation);
- au niveau de l'individu.

CHAPITRE 4

UTILISATION DU SYSTÈME SYCVAR

Les utilisations du système SYCVAR se regroupent en deux parties : la saisie des données et l'exploitation des données.

4.1. SAISIE DES DONNEES

Voyons d'abord comment l'utilisateur peut effectuer l'acquisition des données et quelles sont les facilités mises à sa disposition ?

L'utilisateur doit d'abord déclarer à quelle unité-dossier et à quelle satellité appartiendront les dossiers qu'il désire traiter.

Ceci se fait par les réponses aux questions :

NOM unité-dossier = ? . . .

Satellite = ? . . . ← mnémonique du satellite

Clef personnelle =

Le mode d'affichage des différentes questions du questionnaire est :

<i>N° question - texte en clair - :</i>	<i>Valide</i>
↑	↑
Réponse	Attente de O(oui) ou N(non)

L'entrée proprement dite des informations a été décomposée en 5 phases.

PHASE 1 - Acquisition de l'identification du dossier (n° nom du patient ...)

PHASE 2 - Acquisition du chaînage. Des question sont posées sur le chaînage défini par l'utilisateur.

PHASE 3 - On indique ici le travail que l'on désire faire sur ce dossier par la réponse à la question :

Travail désiré : . Valide : .

Les réponses possibles sont :

1. On déclare entrer un nouveau dossier
2. On désire compléter un dossier dont l'entrée a été abandonnée en cours de route.
3. On désire modifier une ou plusieurs questions d'un dossier.

- Dans le cas où la variable était commandée et ne l'est plus après modification :

Texte Valeur voulez-vous la conserver ? . }_N⁰

- la réponse 0 conserve la valeur ancienne;
 - la réponse N est suivie de la demande de la nouvelle valeur que l'on désire attribuer à cette variable.
- Dans le cas où la variable est commandée efficacement après la modification et la valeur imposée est différente de l'ancienne valeur :
- Texte Valeur Valeur imposée l'acceptez-vous ? . }_N⁰
- la réponse 0 entraîne l'affectation de la valeur imposée à cette variable;
 - la réponse N entraîne l'abandon de la modification dans son ensemble, c'est-à-dire depuis le choix de la variable à modifier.

PHASE 5 - Enregistrement

Toute fin de l'une des fonctions décrites en PHASE 4 est suivie de la question :

Voulez-vous enregistrer ce dossier ? . }_N⁰

La réponse OUI entraîne l'enregistrement définitif

La réponse NON entraîne l'abandon du dossier.

On peut ensuite, soit abandonner le programme d'entrée, soit le début de cette tâche, soit l'entrée d'un dossier du même type de satellite par la réponse à la question :

Suite du travail ? ...

Lors de ces différentes phases, des messages d'erreur peuvent être émis.

PHASES 1, 2 et 3

Il se peut que le travail désiré ne corresponde pas avec l'état du dossier identifié (par exemple : on déclare vouloir entrer un nouveau dossier, et celui-ci existe déjà). A chaque erreur possible, un message est édité.

PHASE 4

Les messages d'erreurs correspondent aux vérifications formelles et de bornes pour chacune des variables.

Par exemple : la réponse 200 à la question : Age ?

entraînera le message : "Valeur hors de la fourchette".

De plus, des messages relatifs aux fonctions de cohérence peuvent être émis si celles-ci ne sont pas satisfaites.

4.2. EXPLOITATION DES DONNEES

Nous allons voir un exemple d'exploitation simple traitant des ensembles de sujets.

L'utilisateur indique au programme le mnémonique associé à l'unité-dossier, puis tape sa clé personnelle. Cette phase d'initialisation s'étant correctement déroulée, le programme demande à l'utilisateur le code de la fonction souhaitée.

Fonction souhaitée :

La réception d'un code correct tel que TRIS, LIST, HIST, MOYE, STAT lance le programme correspondant.

TRIS est un programme permettant de constituer des sous-groupes de dossiers répondant à une formule de sélection.

LIST est un programme ayant pour but de "lister" les réponses attribuées à une série de variables, pour chacun des dossiers appartenant à un ensemble défini par l'utilisateur.

HIST permet le calcul des histogrammes pour une série de variables.

MOYE effectue le calcul des moyennes et écart-types d'une liste de variables, pour un ensemble de dossiers.

Voyons maintenant plus en détail l'emploi des fonctions statistiques (STAT).

Le programme pose d'abord à l'utilisateur le choix de l'étude statistique désirée par la question :

Quelle étude voulez-vous faire ?

L'utilisateur doit répondre par le numéro (1 à 10) de l'étude désirée. La réponse "=" permet l'édition de la liste des études proposées. Cette liste est la suivante :

1. Fréquence, intervalle de confiance
2. Moyenne, Variance, Ecart-types
intervalle de confiance pour la moyenne
3. Moyennes, Variances, Ecart-types
coefficient de corrélation, test d'indépendance
4. Régression linéaire
5. Comparaison de deux fréquences
6. Comparaison de deux populations différentes quant à leurs moyennes (comparaison de moyennes)
7. Comparaison de deux populations différentes quant à leur dispersions (comparaison de variances)

8. Comparaison de deux moyennes dans une population
9. Test d'une fréquence donnée
10. Test d'une moyenne donnée.

Le programme demande alors à l'utilisateur de définir le sous-groupe de dossier sur lequel on va travailler par un numéro qui lui aura été affecté par le programme TRIS.

La question posée est :

Numéro de tri :

Exemple :

Numéro de tri : 12

Garçons

Effectif : 50

Formule : Sexe EQ 1.

Après avoir indiqué le numéro de tri, s'affichent le nom de ce tri, son effectif et la formule ayant permis de l'obtenir.

L'utilisateur doit alors définir les variables sur lesquelles il veut travailler c'est-à-dire les différentes questions apparaissant dans les questionnaires constituant l'unité-dossier. La désignation d'une variable se fait par l'indication du numéro de satellite et du numéro de la variable dans le questionnaire. (La recherche d'une variable nécessite également pour les fiches -suites l'occurrence du satellite concerné).

Le type des variables définies doit correspondre à l'étude demandée :

- type "valeur" pour les moyennes, variances, corrélation, régression
- type logique pour les fréquences.

Dans le cas des fréquences, si le type de la variable n'est pas logique, le programme demande à l'utilisateur de fixer une borne, puis lui fait préciser si la fréquence en question est celle des dossiers présentant une valeur supérieure, ou inférieure à la borne fixée.

Pour ce qui est des études 9 et 10, l'utilisateur indique en plus la valeur théorique (fréquence ou moyenne) à laquelle il veut comparer les fréquences ou moyennes expérimentales.

4.3. FACILITE D'UTILISATION

Le système SYCVAR est donc facile à employer et ne requiert pas la participation active d'un informaticien à l'utilisation. La seule contrainte d'utilisation porte sur le problème d'attention et d'application lors du remplissage des questions posées.

Néanmoins, certaines erreurs peuvent être récupérées. Ainsi, après chaque question, la mention "Valide : .", à laquelle on peut répondre par oui ou non, apparaît. Il s'agit là d'une première sécurité. D'autres moyens de rattrapage sont également prévus. Ainsi, l'emploi de certains caractères spéciaux permettent de parer à certaines erreurs.

Exemple :

- le caractère "*" renvoie à la question immédiatement précédente que l'on peut ainsi modifier.
- le caractère "#" annule la valeur répondue.

o

o o

B - AIDE AU DIAGNOSTIC PAR LA MÉTHODE
DE DÉMONSTRATION AUTOMATIQUE DE
THÉORÈMES

CHAPITRE I

NOTIONS ÉLÉMENTAIRES DE LOGIQUE^(*)

(*) Référence : "Symbolic Logic and Mechanical Theorem Proving" (Academic Press)
CHIN-LIANG CHANG - Richard CHAR TUNG LEE

1.1. LOGIQUE PROPOSITIONNELLE

Cette logique porte, comme son nom l'indique, sur des propositions c'est-à-dire des phrases qui sont soit vraies, soit fausses.

Exemple de proposition : P : Jean a de la fièvre.

- P est appelé un atome (proposition atomique), c'est-à-dire une proposition simple à sujet toujours singulier.
- La connection d'atomes par les opérateurs logiques s'appelle : une proposition composée (ou moléculaire)
- Les opérateurs logiques sont : \sim (négation), \cdot (et), $+$ (ou), \rightarrow (si ... alors) et \leftrightarrow (ssi).
- On appellera une expression bien formée une expression définie comme suit :
 1. un atome est une expression
 2. Si G est une expression $\sim G$ aussi
 3. si G et H sont des expressions, alors $(G \cdot H)$, $(G + H)$, $(G \rightarrow H)$ et $(G \leftrightarrow H)$ aussi
 4. Toutes les expressions sont obtenues en appliquant les règles ci-dessus.
- L'ordre sur les opérateurs logiques existe et est le suivant \leftrightarrow , \rightarrow , \cdot , $+$, \sim .
- On peut évaluer toute expression bien formée en donnant des valeurs de vérité (0 ou 1) aux atomes les composant. Cette évaluation se fait comme suit

G	H	$\sim G$	$G \cdot H$	$G + H$	$G \rightarrow H$	$G \leftrightarrow H$
1	1	0	1	1	1	1
1	0	0	0	1	0	0
0	1	1	0	1	1	0
0	0	1	0	0	1	1

Quelques définitions relatives à la logique propositionnelle

- interprétation : consiste en une assignation de valeur 0 ou 1 à chaque proposition
- valide : une expression est dite valide ssi elle est vraie sous toutes ses interprétations (invalide sinon).
- littéral : un littéral est un atome ou la négation d'un atome.
- forme conjonctive normale : une expression F est dite une "forme conjonctive normale" ssi F est de la forme $F := F_1 \cdot F_2 \dots \cdot F_n$ où $n \geq 1$ et où chaque F_i est une disjonction de littéraux.

Toute expression peut être transformée en forme normale.

En effet, les opérateurs \rightarrow et \leftrightarrow peuvent être transformés en \cdot et $+$.

Ainsi, soient deux propositions P et Q

$P \rightarrow Q$ se transforme en $\sim P + Q$ ou $\sim(P \cdot \sim Q)$

$P \leftrightarrow Q$ se transforme en $(P \rightarrow Q) \cdot (Q \rightarrow P)$.

Ceci sera très utile lors de l'élaboration de fonctions logiques. Ces transformations seront indispensables pour pouvoir appliquer la méthode de DAVIS-PUTNAM.

On emploiera aussi la notion de théorème. Nous allons la définir en deux temps ci-après.

- Etant donné les expressions F_1, \dots, F_n et l'expression G, G est une conséquence logique de F_1, \dots, F_n ssi l'expression $((F_1 \cdot F_2 \cdot \dots \cdot F_n) \rightarrow G)$ est valide.
- Si G est une conséquence logique de F_1, \dots, F_n , l'expression $((F_1 \cdot F_2 \cdot \dots \cdot F_n) \rightarrow G)$ est appelé un théorème.

1.2. LOGIQUE DU PREMIER ORDRE

On est obligé d'aller au-delà de la logique propositionnelle car de nombreuses idées ne peuvent être traitées par ce moyen très simple. En effet, nous devons introduire des notions supplémentaires tels les quantificateurs pour coller à la réalité.

Exemple d'expression logique du 1er ordre

$\forall x$ plus grand que (plus (x,1),x)

où - \forall : quantificateur universel

- x : variable

- plus : fonction (c'est-à-dire une expression dont la signification peut être variable du fait qu'elle mentionne une ou plusieurs variables, mais devient entièrement déterminée si on assigne une signification déterminée à chacune de ses variables (Dopp)).

- plus grand que : prédicats (*)

- 1: constante

On définira une expression bien formée dans la logique du premier ordre récursivement comme suit :

1. un atome est une expression
2. Si F et G sont des expressions alors $\sim F$, $(F \cdot G)$, $(F + G)$, $(F \rightarrow G)$ et $(F \leftrightarrow G)$ aussi.
3. Si F est une expression et x une variable libre dans F (c.à.d. pas à la portée d'un quantificateur de F), alors $(\forall x) F$ et $(\exists x) F$ sont des expressions.
4. Les expressions sont générées par un nombre fini d'applications de (1), (2) et (3).

(*) Prédicat : expression dont la signification est indéterminée à raison du fait qu'elle mentionne des variables, mais devient vraie ou fausse si on assigne une signification déterminée à chacune de ses variables.

Quelques définitions relatives à la logique du 1er ordre

- interprétation : une interprétation d'une expression F dans le premier ordre logique consiste, dans le domaine non vide D en une assignation de valeurs à chaque variable, à chaque fonction et à chaque prédicat de F comme suit :
 - à chaque variable, on assigne un élément de D ;
 - à chaque fonction de n composantes (n -fct), on assigne une projection de D^n sur D ;
 - à chaque prédicat de n -composantes (n -prédicat), on assigne une projection de D^n sur $\{0,1\}$.
- inconsistance : une expression G est inconsistante ssi il n'existe aucune interprétation qui satisfasse G (consistante sinon).
- validité : une expression G est valide ssi toute interprétation satisfait G (invalidé sinon).

Nous prendrons les mêmes définitions que pour la logique propositionnelle en ce qui concerne la conséquence logique et le théorème.

- forme normale Prenex :

But : c'est un prolongement de la forme conjonctive adaptée au premier ordre.

Il en poursuit donc le même but : simplifier et standardiser la forme des expressions afin d'appliquer plus aisément les procédures de démonstration automatique que nous verrons plus loin.

Définition : une expression F dans la logique du premier ordre est dite une forme normale PRENEX ssi l'expression est de la forme :

$$(Q_1 x_1) (Q_2 x_2) \dots (Q_n x_n) (M)$$

où chaque $(Q_i x_i)$ $i = 1, \dots, n$ est un quantificateur soit universel, soit existentiel et M est une expression ne contenant aucun quantificateur.

$(Q_1 x_1) \dots (Q_n x_n)$ est appelé la préfixe.

M est appelé la matrice.

Etapes nécessaires pour arriver à cette forme normale PRENEX

- a. Effectuer les transformations portant sur les opérations \rightarrow et \leftrightarrow

$$- F \leftrightarrow G = (F \rightarrow G) \cdot (G \rightarrow F)$$

$$- F \rightarrow G = \sim F + G$$

- b. Utiliser les propriétés suivantes :

$$- \sim(\sim F) = F$$

$$- \sim(F+G) = \sim F \cdot \sim G$$

$$- \sim(F \cdot G) = \sim F + \sim G$$

$$- \neg(\forall x) F[x] = (\exists x) (\neg F[x])$$

$$- \neg(\exists x) F[x] = (\forall x) (\neg F[x])$$

c. Renommez les variables liées si nécessaire

d. Utiliser les propriétés suivantes :

$$- (Qx) F[x] + G = (Qx) (F[x] + G)$$

$$- (Qx) F[x] \cdot G = (Qx) (F[x] \cdot G)$$

$$- (\forall x) F[x] \cdot (\forall x) H[x] = (\forall x) (F[x] \cdot H[x])$$

$$- (\exists x) F[x] + (\exists x) H[x] = (\exists x) (F[x] + H[x])$$

$$- (Q_1x) F[x] + (Q_2x) H[x] = (Q_1x) (Q_2z) (F[x] + H[z])$$

$$- (Q_3x) F[x] + (Q_4x) H[x] = (Q_3x) (Q_4z) (F[x] \cdot H[z])$$

1.3. THEOREME DE HERBRAND

Principe

Herbrand a développé un algorithme pour trouver une interprétation qui falsifie une expression donnée. Cependant si l'expression est valide, aucune interprétation de ce genre n'existe et l'algorithme stoppera après un nombre fini d'essais.

La procédure ne sera pas appliquée à l'expression elle-même, mais à la négation de l'expression dont on démontre l'inconsistance.

DAVIS et PUTNAM exploitèrent les idées suivantes pour arriver à formaliser la méthode de HERBRAND :

1. Une expression du premier ordre logique peut être transformée en une forme normale PRENEX.
2. La matrice peut être transformée en une forme normale conjonctive.
3. Les quantificateurs existentiels du préfixe peuvent être éliminés par les fonctions SKOLEM.

Que sont les fonctions SKOLEM ?

Soit $F := (Q_1 x_1) (Q_2 x_2) \dots (Q_n x_n) M$

Soit Q_r : quantificateur existentiel, alors

- si aucun quantificateur universel n'apparaît avant Q_r , on choisit une nouvelle constante c et on remplace x_r par c , on supprime aussi $(Q_r x_r)$ du préfixe.
- si $(Q_r x_r)$ apparaît après m quantificateurs universels, on choisit une m -fonction et on remplace x_r par $f(x_1, \dots, x_m)$ on supprime aussi $(Q_r x_r)$ du préfixe.

Exemple

$$(\forall x) (\exists y) (\exists z) ((\neg P(x,y) \cdot Q(x,z)) + R(x,y,z))$$

cette expression se trouve déjà sous forme PRENEX.

- Il nous faut donc la mettre sous forme conjonctive, ce qui donne :

$$(\forall x) (\exists y) (\exists z) ((\neg P(x,y) + R(x,y,z)) \cdot (Q(x,y) + R(x,y,z)))$$

- Passons maintenant aux transformations par les fonctions SKOLEM :

$$(\forall x) ((\neg P(x, f(x)) + R(x, f(x), g(x))) \cdot (Q(x, g(x)) + R(x, f(x), g(x)))) \quad (1)$$

On peut partir de cet exemple pour définir la notion de clause.

Une clause est une disjonction finie de zéro ou plus littéraux.

Ainsi $(\neg P(x, f(x)) + R(x, f(x), g(x)))$ de (1) est une clause.

On peut envisager une autre façon de noter une expression sous la forme PRENEX normale CONJONCTIVE et SKOLEM. On envisage cette expression comme un ensemble dont chaque élément est une clause.

Définition : un ensemble S de clauses est considéré comme la conjonction de toutes les clauses de S, où chaque variable S est sous la coupe d'un quantificateur universel.

Exemple (1) devient :

$$\{ \neg P(x, f(x)) + R(x, f(x), g(x)), Q(x, g(x)) + R(x, f(x), g(x)) \}$$

C'est à partir de cette notion de clauses que nous allons travailler. Pour cela, il faut que nous établissions le passage d'une expression (sous forme conjonctive PRENEX) à un ensemble de clauses. Ceci est réalisé par le théorème suivant.

Théorème

Soit S un ensemble de clauses qui représente la forme normale de l'expression F.

Alors F est inconsistant ssi S est inconsistant.

On sait qu'un ensemble S de clauses est inconsistant ssi il est faux sous toutes ses interprétations sur tous les domaines. Mais il est évidemment impossible de considérer toutes ces interprétations. Heureusement, il existe un domaine H (appelé UNIVERS DE HERBRAND) tel que S est inconsistant ssi S est faux sous toutes les interprétations de ce domaine.

Nous allons maintenant définir ce domaine H.

Définition

- Soit H_0 , l'ensemble des constantes apparaissant dans S . Si aucune constante n'apparaît dans S , alors H_0 consiste en une constante unique, $H_0 = a$
- Pour $i = 0, 1, 2, \dots$, H_{i+1} est l'union des H_i plus l'ensemble de tous les termes de la forme $f^n(t_1, \dots, t_n)$ pour toutes les n -fonctions apparaissant dans S , où les t_j ($j = 1, \dots, n$) sont membres des ensembles H_i .
Ainsi chaque H_i est appelé "l'ensemble des constantes de niveau i " de S et $H = \lim_{i \rightarrow \infty} H_i$ est appelé "l'univers de Herbrand" de S .

Exemple

Soit $S = \{P(f(x), a, g(y), b)\}$

alors $H_0 = \{a, b\}$

$H_1 = \{a, b, f(a), f(b), g(a), g(b)\}$

$H_2 = \{a, b, f(a), f(b), g(a), g(b), f(f(a)), f(f(b)), f(g(a)), f(g(b)), g(g(a)), g(g(b)), g(f(a)), g(f(b))\}$.

Une notion importante et qui reviendra souvent dans la suite est la présence ou l'absence de variables.

Définition

Un "moment sans variable" d'une clause C d'un ensemble S de clauses est une clause obtenue en remplaçant les variables de C par les membres de l'Univers de Herbrand de S .

Exemple :

Soit $S = \{P(x), Q(f(y)) + R(y)\}$

Alors, soit $C = P(x)$ une clause de S et $H = \{a, f(a), f(f(a)) \dots\}$

Ainsi $P(a)$, $P(f(f(a)))$ sont des "moments sans variable" de la clause C .

Nous avons déjà vu la notion d'interprétation tout à fait générale, nous avons vu également la notion d'univers de Herbrand. Il nous faut encore faire le lien entre ces deux notions afin de profiter pleinement de l'utilité de ces deux notions.

C'est ainsi que l'on en vient au concept de H -interprétation c'est-à-dire une interprétation sur l'univers de Herbrand.

Définition

Soient S : un ensemble de clauses

H : univers de Herbrand de S

I : interprétation de S sur H .

Alors I est appelée "H-interprétation" de S s'il satisfait aux conditions suivantes :

1. Il projette toutes les constantes de S sur elles-mêmes.
2. Soit f : n -fonction et h_1, h_2, \dots, h_n , éléments de H .
Dans I , f assigne une fonction qui projette (h_1, \dots, h_n) sur $f(h_1, \dots, h_n)$ qui sont être un élément de H .
3. Pas de restrictions pour les prédicats.

L'utilité de ces concepts est de limiter la définition de l'inconsistance qui est :

"Un ensemble de S de clauses est inconsistant ssi S est faux sous toutes ses interprétations".

Il est évidemment impossible de tester toutes ces interprétations. Nous limitons par le théorème suivant le champs de recherche de ces interprétations en ne travaillant que sur les H-interprétations.

Théorème

Un ensemble S de clauses est inconsistant ssi S est faux sous toutes les H-interprétations de S .

Remarque : on peut également rechercher la preuve pour un ensemble de clauses par la génération d'un arbre sémantique.

Nous n'allons pas nous attarder sur ce point qui est surtout intéressant pour la démonstration du théorème de Herbrand. Nous nous contenterons seulement de citer et d'utiliser le théorème de Herbrand sans le démontrer.

THEOREME DE HERBRAND

Un ensemble S de clauses est inconsistant ssi il existe un ensemble de "moments sans variable" de S qui soit fini et inconsistant.

Exemples

- a. Soit $S = \{P(x), \sim P(f(a))\}$ cet ensemble est inconsistant par le théorème de Herbrand, il existe un ensemble S' de "moments sans variable" de S qui soit fini et inconsistant.

En effet $S' = \{P(f(a)), \sim P(f(a))\}$

- b. Soit $S = \{\neg P(x) + Q(f(x), x), P(g(b)), \neg Q(y, z)\}$
 on applique le même théorème. S inconsistant car
 $S' = \{\neg P(g(b)) + Q(f(g(b)), g(b)), P(g(b)), \neg Q(f(g(b)), g(b))\}$
 est fini et inconsistant.

Le théorème de Herbrand suggère une procédure de "réfutation".

C'est, étant donné un ensemble S de clauses, de prouver son inconsistance par la génération d'ensembles S'_1, S'_2, \dots, S'_n de "moments sans variable" et de tester ces ensembles S'_1, S'_2, \dots, S'_n pour inconsistance.

C'est la méthode de DAVIS et PUTNAM que nous emploierons pour générer ces ensembles que nous allons étudier ci-après.

1.4. METHODE DE DAVIS ET PUTNAM

Nous travaillons sur un ensemble S qui est un ensemble de clauses sans variable.

La méthode consiste en l'application des quatre règles suivantes :

I. Tautologie

- Supprimer de S toutes les clauses (sans variable) qui sont des tautologies.
- L'ensemble S' restant est inconsistant ssi S est inconsistant.

II. Littéral unique

- S'il y a une clause (sans variable) unique L dans S (c'est-à-dire une clause composée d'un seul littéral), on obtient S' de S en supprimant les clauses de S contenant L .
- Si S' est vide, S est consistant (satisfaisable).
- Sinon, on obtient l'ensemble S'' de S' en supprimant $\neg L$ des clauses de S' .
 Si une des clauses S' devient vide par cette suppression, S'' est inconsistant
 S'' est inconsistant ssi S est inconsistant.

III. Littéral pur

- Un littéral L d'une clause (sans variable) de S est dit "pur" dans S ssi le littéral $\neg L$ n'apparaît dans aucune clause de S .
- Si le littéral L est pur dans S , supprimer toutes les clauses contenant L .
- L'ensemble S' restant est inconsistant ssi S est inconsistant.

IV. Partage

- Si l'ensemble S est de la forme :
 $(A_1 + L) \cdot (A_2 + L) \dots (A_m + L) \cdot (B_1 + \neg L) \cdot \dots (B_n + \neg L) \cdot R$
 où A_i, B_i et R sont libres de L et $\neg L$. Alors on obtient les deux ensembles S_1 et S_2 à partir de S avec :

$$S_1 = A_1 \cdot \dots \cdot A_m \cdot R$$

$$S_2 = B_1 \cdot \dots \cdot B_m \cdot R$$

- S est inconsistant ssi $(S_1 + S_2)$ est inconsistant c'est-à-dire si S_1 et S_2 sont tous deux inconsistants.

o

o o

CHAPITRE II

APPLICATION DE LA LOGIQUE AU PROBLÈME MÉDICAL

2.1. UTILISATION DE LA METHODE DE DAVIS ET PUTNAM SUR UN EXEMPLE SIMPLE. CREATION DES ENSEMBLES DE CLAUSES ET EXEMPLE D'UTILISATION.

Supposons que nous ayons un problème simple de diagnostic à résoudre.

Supposons 4 signes a,b,c et d et 2 diagnostics possible D_1 et D_2 .

Supposons encore que sur base de la connaissance médicale, nous avons pu engendrer certaines associations de signes impliquant à coup sûr un diagnostic.

$$\text{Soient } \begin{cases} \neg a \cdot b \cdot \neg c \longrightarrow D_1 \\ \text{diagnostic } D_1 \begin{cases} a \cdot b \cdot d \longrightarrow D_1 \\ b \cdot \neg c \cdot d \longrightarrow D_1 \end{cases} \end{cases}$$

$$\text{diagnostic } D_2 \begin{cases} a \cdot \neg b \cdot d \longrightarrow D_2 \\ a \cdot \neg b \cdot c \longrightarrow D_2 \end{cases}$$

On obtient ainsi deux fonctions L_1 et L_2 entraînant respectivement D_1 et D_2

$$L_1 : (\neg a \cdot b \cdot \neg c) + (a \cdot b \cdot d) + (b \cdot \neg c \cdot d)$$

$$L_2 : (a \cdot \neg b \cdot d) + (a \cdot \neg b \cdot c).$$

Le problème d'aide au diagnostic se pose comme suit :

un patient est représenté par une association (m_j) de signes ou symptômes qui lui sont propres. On va tenter d'affecter un diagnostic à ce malade.

Si on n'y arrive pas, on le place dans une branche de rejet où des examens complémentaires seront nécessaires.

Pour notre exemple simple, on essaie donc de prouver que ($m_j \longrightarrow D_1$).

Si on n'y parvient pas, on essaie de prouver que ($m_j \longrightarrow D_2$).

Si on n'y parvient toujours pas, on se trouve alors dans la branche de rejet.

Pour prouver que ($m_j \longrightarrow D_i$) ($i=1,2$), nous allons employer un moyen indirect en utilisant les fonctions L_i créées plus haut.

En effet, ($m_j \longrightarrow L_i$) est la condition suffisante pour prouver que ($m_j \longrightarrow D_i$) étant donné que ($L_i \longrightarrow D_i$) est toujours vrai vu le choix des associations ayant contribué à la formation des deux fonctions L .

[On utilise la transitivité de " \longrightarrow " ($m_j \longrightarrow L_i$)ⁱ et ($L_i \longrightarrow D_i$) alors ($m_j \longrightarrow D_i$)]
ou encore

$$m_j \longrightarrow L_i \Rightarrow m_j \longrightarrow D_i \text{ car } L_i \longrightarrow D_i \text{ est toujours vérifié}$$

Nous allons utiliser les règles de DAVIS et PUTNAM afin de vérifier que cette application ($m_j \longrightarrow L_i$) est bien un théorème.

Prouver que ($m_j \longrightarrow L_i$) est un théorème revient à prouver que $\sim(m_j \longrightarrow L_i)$ est inconsistent.

Rappel de deux règles de logique

$$F \rightarrow G = \sim F + G = \sim (F \cdot \sim G)$$

$$\sim(\sim F) = F$$

Nous avons vu que pour utiliser les règles de DAVIS et PUTNAM, l'expression $\sim(m_j \rightarrow L_i)$ doit être mise sous forme normale conjonctive et ne doit contenir aucune variable.

L'expression contenant les signes du patient (m_j) et la fonction L_i ne contiennent aucune variable. Il nous reste donc à la mettre sous forme normale conjonctive en utilisant les 2 règles du rappel

$$\sim(m_j \rightarrow L_i) \text{ devient alors } \sim(\sim(m_j \cdot \sim L_i))$$

$$\text{ou encore } (m_j \cdot \sim L_i).$$

Si nous réussissons donc à prouver que $(m_j \cdot \sim L_i)$ est inconsistent, c'est que le patient (m_j) présente toutes les caractéristiques du diagnostic D_i .

Pour tous les patients traités, nous travaillerons toujours sur les mêmes fonctions L_i niées. C'est la raison pour laquelle nous mémorisons ces fonctions sur disque (cfr. supra : utilisation des programmes). Elles représentent la connaissance médicale relative à une pathologie.

Pour cet exemple, les deux fonctions niées L_1 et L_2 seront mémorisées

$$\sim L_1 = (a + \sim b + c) \cdot (\sim a + \sim b + \sim d) \cdot (\sim b + c + \sim d)$$

$$\sim L_2 = (\sim a + b + \sim d) \cdot (\sim a + b + \sim c).$$

Ces fonctions niées peuvent également s'écrire sous forme d'ensembles de clauses (nous les appelons aussi systèmes d'axiomes)

$$\sim L_1 = \{a + \sim b + c, \sim a + \sim b + \sim d, \sim b + c + \sim d\}$$

$$\sim L_2 = \{\sim a + b + \sim d, \sim a + b + \sim c\}$$

Exemple d'utilisation

Supposons une association de signes m_j : $a, b, \sim c, d$

Par la démarche vue plus haut, nous avons vu que : essayer de prouver que :

$(a \cdot b \cdot \sim c \cdot d \rightarrow D_i)$ revient à essayer de prouver que : $(a \cdot b \cdot \sim c \cdot d \cdot \sim L_i)$ est inconsistent ($i=1,2$).

Pour la facilité d'emploi des règles de Davis - Putnam, mettons cette expression sous forme d'ensemble de clauses. Ce qui donne :

pour $i=1$ c'est-à-dire le test par rapport au diagnostic D_1

$$S = \{a, b, \sim c, d, a + \sim b + c, \sim a + \sim b + \sim d, \sim b + c + \sim d\}$$

Soit C l'ensemble de départ mis sous forme aisée pour bien visualiser les transformations

S

- a
- b
- $\sim c$
- d
- $a + \sim b + c$
- $\sim a + \sim b + \sim d$
- $\sim b + c + \sim d$

- a est un littéral unique. On va lui appliquer la règle II de Davis et Putnam. On obtient S_1 en supprimant les clauses contenant a .

S_1

- b
- $\sim c$
- d
- $\sim a + \sim b + \sim d$
- $\sim b + c + \sim d$

S_1 n'est pas vide. On obtient S_2 de S_1 en supprimant $\sim a$ des clauses de S_1

S_2

- b
- $\sim c$
- d
- $\sim b + \sim d$
- $\sim b + c + \sim d$

- b est un littéral unique. On va lui appliquer la règle II. On obtient ainsi :

S_3

- $\sim c$
- d
- $\sim d$
- $c + \sim d$

- $\sim c$ est un littéral unique. On va lui appliquer la règle II. On obtient ainsi :

S_4

- d
- $\sim d$
- $\sim d$

il reste donc ainsi $d \cdot \sim d \cdot \sim d = d \cdot \sim d$. On applique la règle II de Davis et Putnam. On supprime $\sim d$ de l'ensemble des clauses. On obtient ainsi une clause vide par l'application de la 2ème partie de la règle II. S_4 est donc inconsistent D'où l'expression de départ $a \cdot b \cdot \sim c \cdot d \rightarrow D_1$ est un théorème et D_1 est ainsi une conséquence logique des signes présentés.

pour $i = 2$ c'est-à-dire par rapport au diagnostic D_2 , il s'agit exactement de la même démarche. Nous allons donc directement donner la composition de S , l'ensemble des clauses de départ :

$$S = \{a, b, \sim c, d, \sim a + b + \sim d, \sim a + b + \sim c\}$$

Démonstration par Davis et Putnam.

$$\begin{array}{l} S \\ \quad - a \\ \quad - b \\ \quad - \sim c \\ \quad - d \\ \quad - \sim a + b + \sim d \\ \quad - \sim a + b + \sim c \end{array}$$

- a est un littéral unique. On applique la règle II

$$\begin{array}{l} S_1 \\ \quad - b \\ \quad - \sim c \\ \quad - d \\ \quad - b + \sim d \\ \quad - b + \sim c \end{array}$$

- b est un littéral pur. On applique la règle III. On supprime toutes les clauses contenant b

$$\begin{array}{l} S_2 \\ \quad - \sim c \\ \quad - d \end{array}$$

- $\sim c$ et d étant deux littéraux uniques, on les supprime. L'ensemble restant est vide d'où S est consistant (satisfaisable).
L'expression de départ c'est-à-dire $(a \cdot b \cdot \sim c \cdot d \rightarrow D_2)$ est donc non valide et ne forme pas un théorème.

2.2. UTILISATION DE LA METHODE DE DAVIS ET PUTNAM SUR LA PATHOLOGIE THYROIDIENNE CREATION DES ENSEMBLES DE CLAUSES ET EXEMPLE D'UTILISATION

Nous allons utiliser la méthode de DAVIS et PUTNAM de façon complètement analogue à celle vue au point 2.1. sur un exemple simple de diagnostic.

Nous avons déjà étudié lors de l'application booléenne du diagnostic la manière dont ont été obtenus les signes les plus discriminants de cette pathologie, ainsi que les associations de ces signes formant la connaissance médicale.

Nous avons vu que sur base de ces associations, quatre fonctions logiques ont été construites :

- L_1 impliquant le diagnostic de l'hypothyroïde
- L_2 impliquant le diagnostic d'état normal
- L_3 impliquant le diagnostic d'hyperthyroïde
- L_4 impliquant le diagnostic de thyréoses.

Pour une question de facilité et de rapidité d'écriture, nous avons remplacé les signes des patients par des lettres S suivies d'un numéro. Ceci est rendu aisé par le nombre réduit de signes retenus (pas plus de 20 signes). Pour trouver le signe du patient ou le numéro d'un signe, il suffit de consulter la table où se trouvent tous les signes considérés avec un S numéroté.

Remarque : si la méthode de Davis et Putnam devenait opératoire dans un hôpital, on représenterait les variables considérées par des mnémoniques avec des tables de synonymes.

S_1	Frilosité
S_2	Asthénie après effort
S_3	Peau sèche et squameuse
S_4	Augmentation de poids
S_5	Taux de fixation 6ème heure diminué
S_6	Taux de fixation 24ème heure diminué
S_7	Test Blanquet - Meyniel III diminué
S_8	Réflexogramme VP > 200 ms
S_9	Voix rauque et nasonnée
S_{10}	Blanquet - Meyniel III augmenté
S_{11}	Taux de fixation 24ème heure augmenté
S_{12}	Tachycardie
S_{13}	Amaigrissement
S_{14}	Thermophobie
S_{15}	Soif
S_{16}	Taux de fixation 6ème heure augmenté
S_{17}	Taux de fixation 24ème heure augmenté

Tableau de correspondance des signes

A partir de cette table des signes et des fonctions booléennes telles qu'elles ont été énoncées auparavant, nous pouvons écrire les quatre fonctions suivantes :

$$L_1 : (S_1+S_2) \cdot (S_3+S_4) \cdot (S_5+S_6) \cdot S_7 \cdot (S_8+S_9) \quad [\text{Hypothyroïde}]$$

$$L_2 : -S_{10} \cdot -S_7 \cdot -S_{11} \cdot -S_6 \quad [\text{Normal}]$$

$$L_3 : (S_{12}+S_{13}) \cdot (S_{14}+S_{15}) \cdot (S_{16}+S_{11}) \cdot S_{10} \quad [\text{Hyperthyroïde}]$$

$$L_4 : S_{11} \cdot S_{17} \cdot S_{10} \cdot (-S_{12}+-S_{13}) \cdot (-S_{14}+-S_{15}) \quad [\text{Thyréoses}]$$

Nous avons énoncé, lors du point précédent, le problème de l'aide au diagnostic. Nous en sommes arrivés aux conclusions suivantes.

Soit m_j : association de signes présentés par un nouveau patient

D_i : ($i = 1,4$) quatre diagnostics de groupe de la thyroïde

L_i : ($i = 1,4$) quatre fonctions logiques telles que ($L_i \rightarrow D_i$)

Prouver que ($m_j \rightarrow D_i$) est un théorème revient à prouver que ($m_j \cdot \sim L_i$) est inconsistant (cfr. exemple simple de DAVIS et PUTNAM du point 2.1.).

Pour pouvoir utiliser Davis et Putnam afin de démontrer l'inconsistance de l'expression obtenue, cette dernière doit se trouver sous forme normale conjonctive. Or, les (L_i) ne se présentent pas sous cette forme et doivent donc subir quelques transformations avant d'être utilisés.

C'est pourquoi nous allons développer les quatre fonctions par la propriété de distributivité afin d'obtenir ces fonctions sous forme canonique.

Propriété utilisée :

- $a \cdot (b + c) = (a \cdot b) + (a \cdot c)$
- $(a + b) \cdot (c + d) = (a \cdot c) + (a \cdot d) + (b \cdot c) + (b \cdot d)$

Transformations

$$\begin{aligned} L_1 &= (S_1+S_2) \cdot (S_3+S_4) \cdot (S_5+S_6) \cdot S_7 \cdot (S_8+S_9) \\ &= [(S_1 \cdot S_3) + (S_1 \cdot S_4) + (S_2 \cdot S_3) + (S_2 \cdot S_4)] \cdot (S_5+S_6) \cdot S_7 \cdot (S_8+S_9) \\ &= [(S_1 \cdot S_3 \cdot S_5) + (S_1 \cdot S_4 \cdot S_5) + (S_2 \cdot S_3 \cdot S_5) + (S_2 \cdot S_4 \cdot S_5) + (S_1 \cdot S_3 \cdot S_6) + \\ &\quad (S_1 \cdot S_4 \cdot S_6) + (S_2 \cdot S_3 \cdot S_6) + (S_2 \cdot S_4 \cdot S_6)] \cdot S_7 \cdot (S_8+S_9) \\ &= (S_1 \cdot S_3 \cdot S_5 \cdot S_7 \cdot S_8) + (S_1 \cdot S_4 \cdot S_5 \cdot S_7 \cdot S_8) + (S_2 \cdot S_3 \cdot S_5 \cdot S_7 \cdot S_8) + \\ &\quad (S_2 \cdot S_4 \cdot S_5 \cdot S_7 \cdot S_8) + (S_1 \cdot S_3 \cdot S_6 \cdot S_7 \cdot S_8) + (S_1 \cdot S_4 \cdot S_6 \cdot S_7 \cdot S_8) + \\ &\quad (S_2 \cdot S_3 \cdot S_6 \cdot S_7 \cdot S_8) + (S_2 \cdot S_4 \cdot S_6 \cdot S_7 \cdot S_8) + (S_1 \cdot S_3 \cdot S_5 \cdot S_7 \cdot S_9) + \\ &\quad (S_1 \cdot S_4 \cdot S_5 \cdot S_7 \cdot S_9) + (S_2 \cdot S_3 \cdot S_5 \cdot S_7 \cdot S_9) + (S_2 \cdot S_4 \cdot S_5 \cdot S_7 \cdot S_9) + \\ &\quad (S_1 \cdot S_3 \cdot S_6 \cdot S_7 \cdot S_9) + (S_1 \cdot S_4 \cdot S_6 \cdot S_7 \cdot S_9) + (S_2 \cdot S_3 \cdot S_6 \cdot S_7 \cdot S_9) + \\ &\quad (S_2 \cdot S_4 \cdot S_6 \cdot S_7 \cdot S_9) \end{aligned}$$

L_2 n'est pas modifiée $-S_{10} \cdot -S_7 \cdot -S_{11} \cdot -S_6$

$$\begin{aligned} L_3 &= (S_{12}+S_{13}) \cdot (S_{14}+S_{15}) \cdot (S_{16}+S_{11}) \cdot S_{10} \\ &= (S_{12} \cdot S_{14} \cdot S_{16} \cdot S_{10}) + (S_{12} \cdot S_{14} \cdot S_{11} \cdot S_{10}) + (S_{12} \cdot S_{15} \cdot S_{16} \cdot S_{10}) \\ &\quad + (S_{12} \cdot S_{15} \cdot S_{11} \cdot S_{10}) + (S_{13} \cdot S_{14} \cdot S_{16} \cdot S_{10}) + (S_{13} \cdot S_{14} \cdot S_{11} \cdot S_{10}) \\ &\quad + (S_{13} \cdot S_{15} \cdot S_{16} \cdot S_{10}) + (S_{13} \cdot S_{15} \cdot S_{11} \cdot S_{10}) \end{aligned}$$

$$\begin{aligned} L_4 &= S_{11} \cdot S_{17} \cdot S_{10} \cdot (-S_{12}+-S_{13}) \cdot (-S_{14}+-S_{15}) \\ &= (S_{11} \cdot S_{17} \cdot S_{10} \cdot -S_{12} \cdot -S_{14}) + (S_{11} \cdot S_{17} \cdot S_{10} \cdot -S_{12} \cdot -S_{15}) + \\ &\quad (S_{11} \cdot S_{17} \cdot S_{10} \cdot -S_{13} \cdot -S_{14}) + (S_{11} \cdot S_{17} \cdot S_{10} \cdot -S_{13} \cdot -S_{15}) \end{aligned}$$

Mais nous avons vu que les fonctions L_i ($i = 1,4$) intervenaient de façon niée ($\sim L_i$).

Propriété utilisée

$$\sim (\sim a \cdot b) = a + \sim b$$

$$\sim (\sim a + b) = a \cdot \sim b$$

Transformations

$$\begin{aligned} (\sim L_1) &= (-S_1+-S_3+-S_5+-S_7+-S_8) \cdot (-S_1+-S_4+-S_5+-S_6+-S_8) \cdot (-S_2+-S_3+-S_5+-S_7+-S_8) \cdot \\ &\quad (-S_2+-S_4+-S_5+-S_7+-S_8) \cdot (-S_1+-S_3+-S_6+-S_7+-S_8) \cdot (-S_1+-S_4+-S_6+-S_7+-S_8) \cdot \\ &\quad (-S_2+-S_3+-S_5+-S_7+-S_8) \cdot (-S_2+-S_4+-S_6+-S_7+-S_8) \cdot (-S_1+-S_3+-S_5+-S_7+-S_9) \cdot \\ &\quad (-S_1+-S_4+-S_5+-S_7+-S_9) \cdot (-S_2+-S_3+-S_5+-S_7+-S_9) \cdot (-S_2+-S_4+-S_5+-S_7+-S_9) \cdot \\ &\quad (-S_1+-S_3+-S_6+-S_7+-S_9) \cdot (-S_1+-S_4+-S_6+-S_7+-S_9) \cdot (-S_2+-S_3+-S_6+-S_7+-S_9) \cdot \\ &\quad (-S_2+-S_4+-S_6+-S_7+-S_9) \end{aligned}$$

$\sim L_2$ ne contient qu'une seule clause

$$\sim L_2 = (S_{10}+S_7+S_{11}+S_6)$$

$$\begin{aligned} \sim L_3 &= (-S_{12}+-S_{14}+-S_{16}+-S_{10}) \cdot (-S_{12}+-S_{14}+-S_{11}+-S_{10}) \cdot (-S_{12}+-S_{15}+-S_{16}+-S_{10}) \cdot \\ &\quad (-S_{12}+-S_{15}+-S_{11}+-S_{10}) \cdot (-S_{13}+-S_{14}+-S_{16}+-S_{10}) \cdot (-S_{13}+-S_{14}+-S_{11}+-S_{10}) \cdot \\ &\quad (-S_{13}+-S_{15}+-S_{16}+-S_{10}) \cdot (-S_{13}+-S_{15}+-S_{11}+-S_{10}) \cdot \end{aligned}$$

$$\begin{aligned} \sim L_4 &= (-S_{11}+-S_{17}+-S_{10}+S_{12}+S_{14}) \cdot (-S_{11}+-S_{17}+-S_{10}+S_{12}+S_{15}) \\ &\quad (-S_{11}+-S_{17}+-S_{10}+S_{13}+S_{14}) \cdot (-S_{11}+-S_{17}+-S_{10}+S_{13}+S_{15}) \end{aligned}$$

Les nouvelles fonctions L_i ainsi obtenues se présentent sous forme normale conjonctive et elles ne contiennent pas de variable.

On peut donc utiliser DAVIS et PUTNAM pour essayer de prouver l'inconsistance de l'expression $(m_j \cdot \sim L_i)$.

Comme application de DAVIS et PUTNAM pour la pathologie thyroïdienne, nous avons développé deux programmes pour la mise en fonction de cette méthode.

C'est ce dont nous parlons au point 2.3.

2.3. ETUDE DES PROGRAMMES CALQUANT LA METHODE DE DAVIS ET PUTNAM

Maintenant que nous avons vu la méthode de Davis et Putnam et son utilisation, nous pouvons étudier les deux programmes utilisés pour calquer cette théorie. Il s'agit des programmes ENTRAX et DEMONS écrits en FORTRAN.

ENTRAX : programme d'entrée des axiomes définissant la connaissance médicale

DEMONS : programme d'acquisition des signes à démontrer et de la démonstration proprement dite.

2.3.1. ENTRAX

2.3.1.1. Axiomes à définir

Il était initialement prévu de créer un programme qui transformerait directement la connaissance médicale sous forme d'axiomes c'est-à-dire de tirer du dossier médical un ensemble de clauses (base de la démonstration). Le temps ne nous a pas permis de terminer ce travail. La conception d'un tel programme est néanmoins simple. En effet, des programmes avaient déjà été construits pour le choix des signes discriminants et pour l'obtention d'associations de ces signes impliquant à coup sûr un diagnostic (cfr. application de la logique booléenne).

A partir de ces associations, il est aisé de trouver les systèmes d'axiomes. En effet, en prenant toutes les associations relatives à un diagnostic $(m_j \rightarrow D_i)$ avec $j = 1$ à n , il suffit de faire un "ou logique" entre ces associations; ce qui donne :

$$(m_1 \rightarrow D_i) + (m_2 \rightarrow D_i) \dots + (m_n \rightarrow D_i)$$

ou encore

$$(m_1 + m_2 + \dots + m_n \rightarrow D_i)$$

où $m_1 + m_2 + \dots + m_n = L$ (fonction logique)

où m_j est un minterm.

Pour obtenir l'ensemble d'axiomes à mémoriser, nous avons vu qu'il fallait nier L, ce qui donne comme ensemble de clauses

$$\{\sim m_1, \sim m_2, \dots, \sim m_n\}.$$

On remarque que tout ceci peut être obtenu de manière automatique.

Cela n'a pas été réalisé, ce qui nous a valu de faire le développement manuel des fonctions thyroïdiennes du point 2.

Nous avons déjà signalé au point précédent que la connaissance médicale restait stable durant un certain temps. C'est pourquoi nous la mémorisons sur disque. C'est cette mémorisation des axiomes ainsi que leur mise à jour que nous allons étudier ci-après.

2.3.1.2. Organigrammes et spécifications sur ENTRAX

ENTRAX est un programme prévu pour l'entrée et la mise à jour des programmes définissant la connaissance médicale en interactif.

Nous allons d'abord étudier les fichiers utilisés par ENTRAX et ensuite le mode d'emploi et les possibilités offertes par ce programme.

a. Utilisation des fichiers

b. Explication de l'organigramme

c. Application simple.

a. Utilisation des fichiers

Les programmes mis au point travaillent avec deux fichiers.

A.1. FICHIER CATDEM

Utilité : CATDEM sert de catalogue où sont répertoriés tous les groupes d'axiomes (toutes les pathologies traitées). On se positionnera donc dans CATDEM à la pathologie choisie pour en connaître les différents diagnostics possibles.

Organisation : ce fichier contient jusqu'à 100 enregistrements du type suivant :

nom du groupe	nombre de systèmes	nom du support disque	nom fichier LIT-CL 1	nom fichier LIT-CL 2	nom fichier LIT-CL 3
8c	2c	6c	6c	6c	6c

Les huit premiers caractères représentent le nom du groupe d'axiomes considéré. A chaque groupe considéré correspond une pathologie donnée (ex. Pathologie Thyroïdienne).

Les deux caractères suivants sont relatifs au nombre de systèmes compris dans le groupe. Chaque système équivaut à un diagnostic. Ces deux caractères vont donc supporter le nombre de diagnostics possibles pour une pathologie (de 0 à 40).

Les six caractères suivants donnent le nom du support disque des fichiers LIT-CL que nous définissons ci-après.

Ensuite, on a 40 fois 6 caractères qui supportent les noms de fichiers LIT et CL qui sont les fichiers de littéraux et de clauses relatifs aux différents diagnostics. On choisit des noms de fichiers évoquant les noms des diagnostics.

Exemple d'enregistrement de CATDEM

THYROIDE	4	BIMS 01	HYP0	NORMAL	HYPER	THYRE0
----------	---	---------	------	--------	-------	--------

4 diagnostics sont répertoriés pour la pathologie thyroïdienne :

- fichier HYP0 : hypothyroïde
- fichier NORMAL : fonctions thyroïdiennes normales
- fichier HYPER : hyperthyroïde
- fichier THYRE0 : thyréoses.

A.2. FICHIERS LIT-CL

Utilité : les fichiers de littéraux et de clauses répertoriés dans CATDEM représentent la connaissance médicale (diagnostics) mise sous forme d'ensembles de clauses. On emploiera parfois le terme "systèmes d'axiomes". Ces tables sont élaborées de la façon décrite ci-après pour une application aisée de DAVIS et PUTNAM.

Organisation

- Table de littéraux : LIT (20,47) c'est-à-dire on peut avoir un maximum de 20 littéraux et de 40 clauses. Chaque ligne correspond à la description d'un littéral de la façon suivante :
 - LIT (i,j) j = 1,6 identificateur du littéral
 - LIT (i,j) j = 7 nombre de clauses dans lesquelles on retrouve le littéral i
 - LIT (i,j) j = 8,47 n° de clause dans lequel on trouve ce littéral. Ce numéro est positif si le littéral n'est pas nié et négatif s'il est nié.
- Table de clauses : CL (40,21). Chaque ligne correspond à l'identification d'une clause de la façon suivante :
 - CL (i,j) j = 1 nombre de littéraux dans la clause
 - CL (i,j) j = 2,21 n° des littéraux intervenant dans la clause.

Le fichier comprend également un vecteur NB(7) contenant différentes caractéristiques relatives à ces deux tables.

Exemple simple

Soient 2 diagnostics : D_1 et D_2

Soient 4 signes : a, b, c et d.

$$\begin{array}{l} \underline{D_1} : \quad \neg a \cdot b \cdot \neg c \longrightarrow D_1 \\ \quad \neg a \cdot b \cdot d \longrightarrow D_1 \\ \quad b \cdot \neg c \cdot d \longrightarrow D_1 \end{array}$$

$$\begin{array}{l} \underline{D_2} : \quad a \cdot \neg b \cdot d \longrightarrow D_2 \\ \quad a \cdot \neg b \cdot c \longrightarrow D_2 \end{array}$$

Ceci nous donne donc 2 systèmes d'axiomes

$$D_1 : (a + \neg b + c) \cdot (a + \neg b + \neg d) \cdot (\neg b + c + \neg d)$$

$$D_2 : (\neg a + b + \neg d) \cdot (\neg a + b + \neg c)$$

Ceci se représente de la façon suivante sur LIT et CL contenant les fonctions niées relatives à ces diagnostics.

D₁Table des littéraux

*	A	*	2	*	1	2	0	0	0	0	...
*	B	*	3	*	-1	-2	-3	0	0	0	...
*	C	*	2	*	1	3	0	0	0	0	...
*	D	*	2	*	-2	-3	0	0	0	0	...

Table des clauses

3	*	1	*	-2	*	3	*	0	*	0	...
3	*	1	*	-2	*	-4	*	0	*	0	...
3	*	-2	*	3	*	-4	*	0	*	0	...

D₂Table des littéraux

*	A	*	2	*	-1	-2	0	0	0	0	...
*	B	*	2	*	1	2	0	0	0	0	...
*	C	*	1	*	-2	0	0	0	0	0	...
*	D	*	1	*	-1	0	0	0	0	0	...

Table des clauses

*	3	*	-1	*	2	*	-4	*	0	*	0	*	...
*	3	*	-1	*	2	*	-3	*	0	*	0	*	...

b. Explication de l'organigramme

Ce programme commence par ouvrir le fichier CATDEM. Après cette ouverture de fichier, le programme demande sur la visu le nom du groupe d'axiomes, c'est-à-dire le nom de la pathologie à traiter.

Ensuite, il demande s'il s'agit d'une option création de nouveaux groupes (paramètre = 1), de mise à jour de groupe existant (=2), de mise à jour avec remise à zéro des informations précédentes (=3) ou d'annulation d'un groupe d'axiomes (=4).

1° Branche de création

Elle appelle, tout d'abord, le sous-programme RGPAX1 qui recherche l'existence dans les fichiers CATDEM d'un enregistrement correspondant au nom de groupe de système lu sur la visu. Le nom qui a été lu ne doit pas se trouver dans le fichier CATDEM. S'il s'y trouve, c'est qu'il existait déjà et un message est affiché.

Le sous-programme RGPAX1, dans le cas où il ne trouve pas un groupe d'axiomes du nom donné, rend dans un de ses paramètres, l'adresse d'un secteur libre dans CATDEM où l'on peut créer le nouveau groupe.

Le sous-programme CREAX1 va donc être appelé pour remplir cet enregistrement avec le nom du groupe du système d'axiomes, le nom du disque supportant les fichiers qui sera demandé sur la visu, ou implicitement repris sur les précédents enregistrements de CATDEM. Il met le nombre de groupes à zéro, met la fin de l'enregistrement à zéro.

Puis cette branche de programme va rejoindre une branche commune à tous les traitements (sauf l'annulation).

2. Branche de traitement : mise à jour sans remise à zéro.

Cette branche appelle le programme RGPAX1, vérifie que le nom du groupe du systèmes d'axiomes existe bien et rejoint la branche commune.

3. Branche de traitement : mise à jour avec remise à zéro

Cette branche appelle successivement le sous-programme RGPAX1, (même raison qu'au 2.) puis le sous-programme SUPAX1 qui est un programme d'annulation du groupe de systèmes d'axiomes et enfin le sous-programme de création CREAX1.

Il rejoint ensuite la branche commune.

4. Branche de traitement : annulation

Cette branche se contente d'appeler le sous-programme SUPAX1 qui remet à zéro l'enregistrement correspondant de CATDEM et efface les fichiers correspondants, puis retourne vers le module d'interrogation visu pour demander si le traitement doit être poursuivi ou si le travail est terminé.

La branche commune poursuit le traitement en demandant la fonction à exécuter, non plus au niveau des groupes de systèmes d'axiomes mais au niveau de chaque système d'axiomes.

Le paramètre lu, pour le traitement de ce niveau, possède lui aussi 4 moments.

1° Branche - création d'un nouveau système d'axiomes

Cette branche appelle le sous-programme REMPL1 qui recherche un emplacement libre parmi les 40 emplacements possibles dans un enregistrement de CATDEM afin d'y noter le nom du fichier contenant les clauses et les littéraux que l'on va être amené à enregistrer.

Le nom du fichier est choisi le plus explicite possible
ex. : HYPO pour le système relatif au diagnostic d'hypothyroïde.
 Cette branche appelle ensuite le sous-programme DEFI00, qui crée effectivement le fichier qui va recevoir les tables des littéraux et des clauses. Cette branche va, ensuite rejoindre une branche commune.

Les 2°, 3° et 4° branches du programme correspondant au traitement de mise à jour avec remise à zéro et effacement du système d'axiomes, passent par un tronc commun qui demande à la visu quel est le numéro du groupe d'axiomes à traiter.

Si le nombre est égal à zéro, le sous-programme CHOIS1 permet un défilement de tous les noms de fichier LIT-CL permettant de repérer le numéro souhaité.

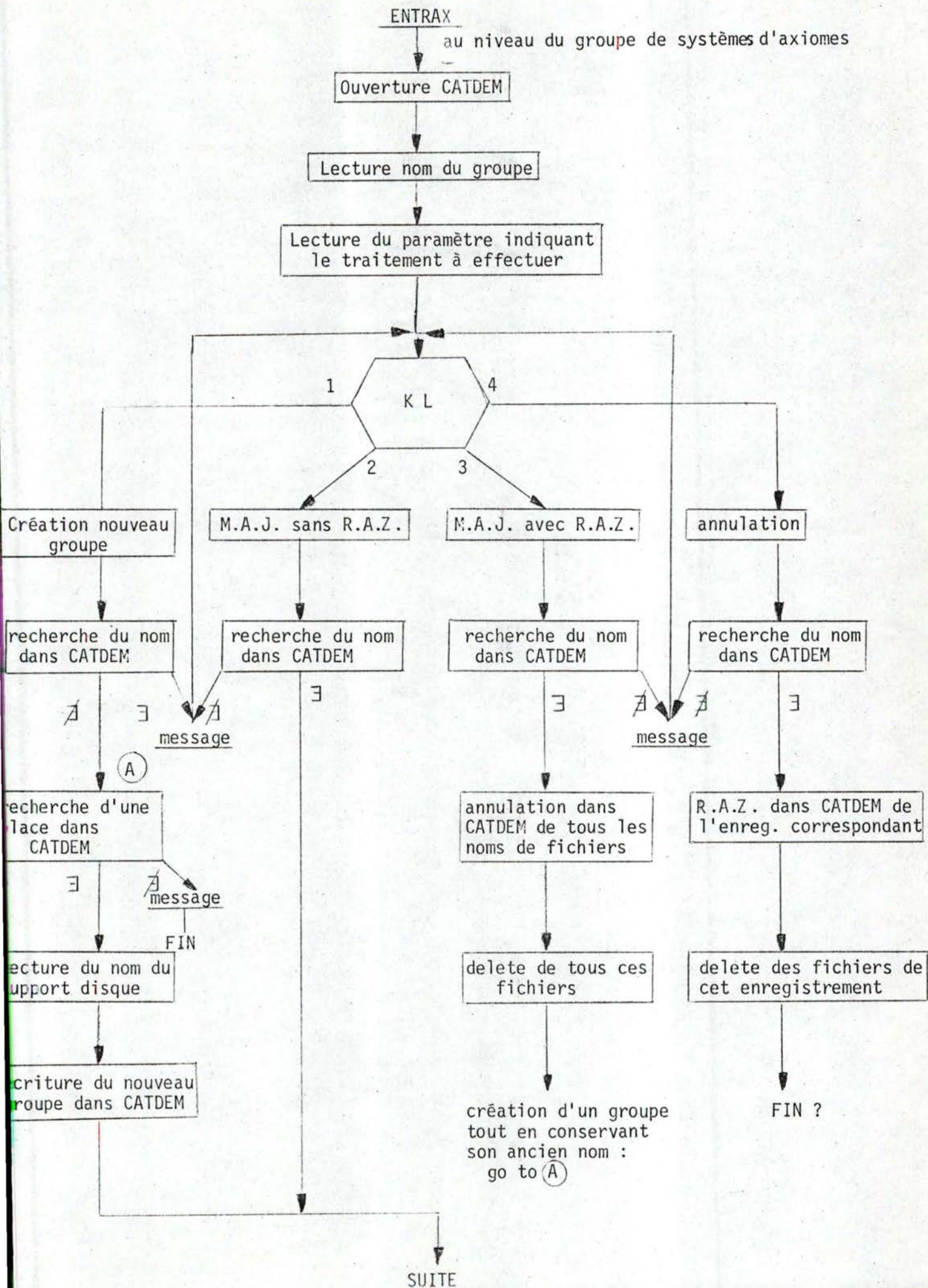
Si l'option de traitement est égal à 2, le programme effectue une lecture préalable des axiomes existants dans le fichier qui vient d'être repéré.

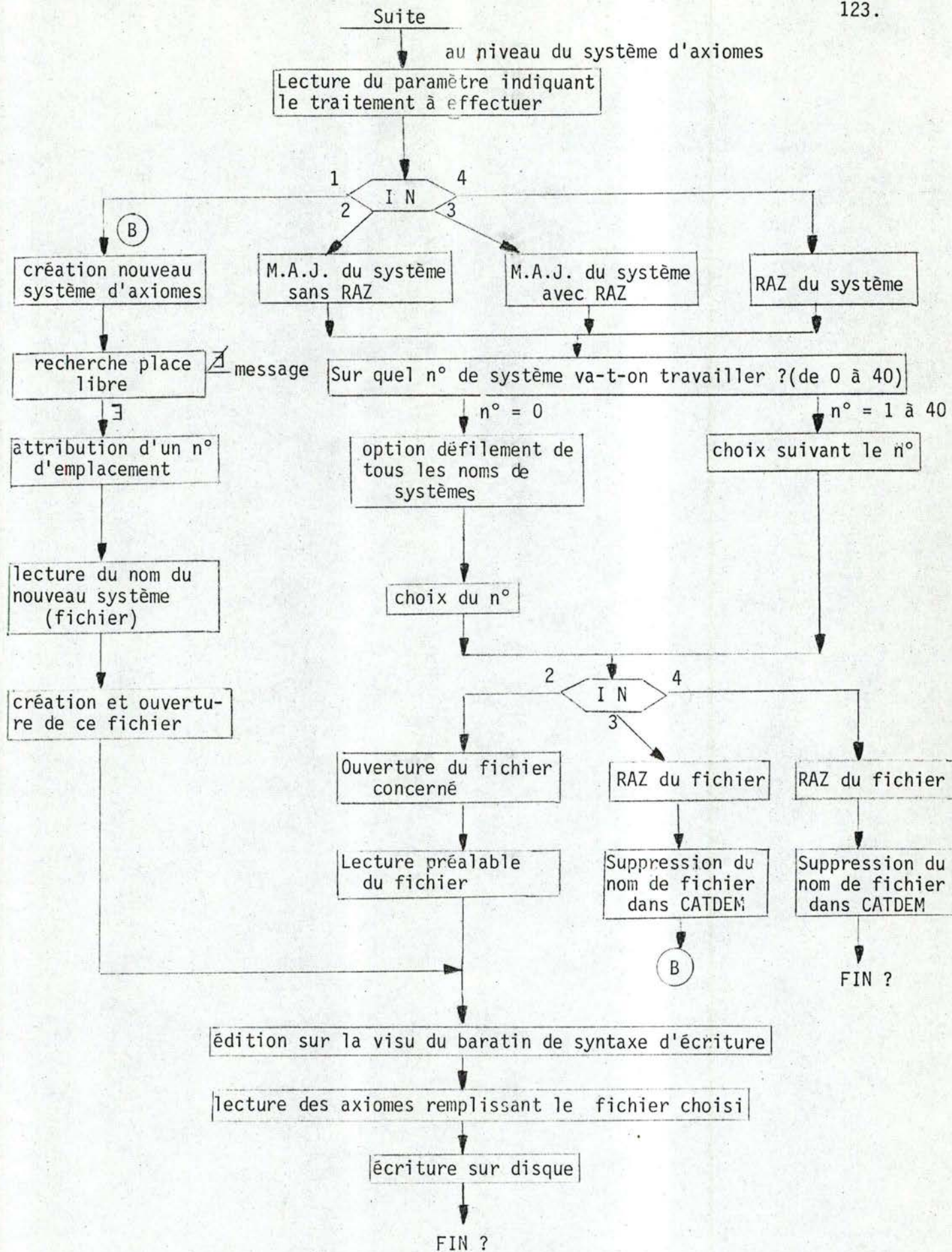
Si l'option mise à jour avec remise à zéro ou effacement d'un système d'axiomes a été demandée, on effectue un appel au programme EFSYS1 qui est chargé d'effacer les coordonnées du système d'axiomes dans CATDEM et efface le fichier LIT-CL correspondant.

Si le traitement demandé était uniquement l'annulation, on retourne au module d'interrogation de la visu pour savoir si on poursuit ou non le traitement; sinon, on retourne en début de la branche de création d'un nouveau système d'axiomes, après avoir mémorisé l'ancien nom du fichier.

Les branches citées, sauf celles d'annulation du système d'axiomes, se poursuivent par une branche commune unique qui effectue l'OPEN du fichier des clauses et littéraux que l'on va créer ou mettre à jour.

Puis, le programme appelle le sous-programme BARAT1 qui édite sur la visu les modalités de la syntaxe des axiomes; enfin, on appelle le sous-programme LANGEN qui procède à la fois à l'analyse des axiomes lors de leur acquisition sur la visu et surtout à la génération des tables correspondantes : littéraux et clauses.





Enfin, ce programme écrit les tables ainsi constituées sur disque avant de retourner au module de poursuite de traitement qui permet de boucler, si l'utilisateur veut continuer sur un nouveau traitement à partir de la lecture du nom du groupe des systèmes d'axiomes sur la visu.

c. Application simple : exemple d'utilisation de ENTRAX

Comme on l'a vu, ENTRAX permet d'introduire de nouveaux groupes de systèmes d'axiomes (c'est-à-dire de nouvelles pathologies) et de nouveaux ensembles de clauses (c'est-à-dire de nouveaux diagnostics).

ENTRAX permet également de les supprimer et de les mettre à jour.

Ce sont des exemples de ces diverses opérations que nous montrons ci-après :

Exemple A du listing

On travaille sur le groupe AXIOME que l'on essaie de créer.

Or ce groupe existait déjà. Un message est alors affiché.

Exemple B du listing

On met à jour le groupe AXIOME en faisant une remise à zéro avec modification de l'un des ensembles de clauses. Ce cas se présente lorsque des signes ou des associations plus discriminantes ont été découverts par les médecins et doivent être pris en compte pour les diagnostics ultérieurs.

L'opératrice choisit le système d'axiomes (le diagnostic) à modifier et introduit ensuite le nouveau système plus performant.

Lors de l'introduction des signes, les clauses (axiomes) sont séparées par le caractère "/"(.), et la dernière clause est terminée par le caractère "*".

Les signes repris dans les différentes clauses sont séparés par un point "." (+).

Exemple C du listing

Cet exemple montre la suppression pure et simple d'un ensemble de clauses. (Un diagnostic).

Ce cas peut se présenter lorsque les résultats relatifs à un diagnostic ne s'avèrent pas assez performants.

On préfère éliminer ce système entraînant des erreurs, quitte à grossir la branche de rejet par les nouveaux cas non diagnostiqués.

Exemple D du listing

Cet exemple montre la suppression de tout un groupe c'est-à-dire de toute une pathologie.

Cela se produit si la connaissance médicale relative à une pathologie est tombée en désuétude ou si de nouveaux signes ou examens la bouleversent totalement.

Exemples E et F du listing

Ces exemples nous montrent la création d'un nouveau groupe comprenant deux systèmes d'axiomes c'est-à-dire deux diagnostics. La pathologie imaginaire créée ici est celle de l'exemple simple de diagnostic vu au point 2.1. de ce chapitre.

ENTRAX laisse donc à l'utilisateur un grand choix de manoeuvres correspondant aux besoins du problème. Néanmoins, l'utilité primordiale de ENTRAX reste la création des groupes de systèmes d'axiomes.

E AXIOME
 CREATION=1,MISE A JOUR=2,MISE A.J.AVEC RAZ=3,ANNULATION=4 .1
 GROUPE DEJA EXISTANT
 NOM DU SYST.? =AXIOME

3. CREATION=1,MISE A JOUR=2,MISE A.J.AVEC RAZ=3,ANNULATION=4 .2
 QUE FAIRE SUR CE GROUPE
 CREATION=1,MISE A JOUR=2,MISE A.J.AVEC RAZ=3,ANNULATION=4 .3
 QUEL NUM.DANS LE GROUPE? ..00
 1 TETE
 2 PIEDS
 NUM.CHOISI?= ..2
 SYST.D'AXIOMES CHOISIS= PIEDS
 NOM DU SYSTEME D AXIOMES A CREER:PIEDS
 SYNTAXE D'ECRITURE DES AXIOMES-C'EST A DIRE DES AFFIRMAIONS DE BASES

LES AXIOMES SONT DES SUITES DE SIGNES,OU DE SIGNES NIES (PRECEDES DU SYMBOLE
 LES AXIOMES SONT SEPARES PAR LE CARACTERE /
 LE DERNIER AXIOME EST TERMINE PAR LE CAR. *
 EXMPLE: TOUX.FIEVRE.CRACHA.RALES/HILEOP.TRAMEA.FIEVRE*
 .A.B.-C.-D*
 NOUVEAU TRT(=0) OU FIN? ..0

C. NOM DU SYST.? =AXIOME
 CREATION=1,MISE A JOUR=2,MISE A.J.AVEC RAZ=3,ANNULATION=4 .2
 QUE FAIRE SUR CE GROUPE
 CREATION=1,MISE A JOUR=2,MISE A.J.AVEC RAZ=3,ANNULATION=4 .4
 QUEL NUM.DANS LE GROUPE? ..2
 SYST.D'AXIOMES CHOISIS= PIEDS
 NOUVEAU TRT(=0) OU FIN? ..0

D. NOM DU SYST.? =AXIOME
 CREATION=1,MISE A JOUR=2,MISE A.J.AVEC RAZ=3,ANNULATION=4 .4
 NOUVEAU TRT(=0) OU FIN? ..0

E. NOM DU SYST.? =AXIOME
 CREATION=1,MISE A JOUR=2,MISE A.J.AVEC RAZ=3,ANNULATION=4 .1
 NOM DISQUE SUPORT SYST.AXION.BINS01
 NUM.ATTRIBUE AU GROUPE= 29
 QUE FAIRE SUR CE GROUPE
 CREATION=1,MISE A JOUR=2,MISE A.J.AVEC RAZ=3,ANNULATION=4 .1
 NOM DU SYSTEME D AXIOMES A CREER:GAUCHE
 SYNTAXE D'ECRITURE DES AXIOMES-C'EST A DIRE DES AFFIRMAIONS DE BASES

LES AXIOMES SONT DES SUITES DE SIGNES,OU DE SIGNES NIES (PRECEDES DU SYMBOLE
 LES AXIOMES SONT SEPARES PAR LE CARACTERE /
 LE DERNIER AXIOME EST TERMINE PAR LE CAR. *
 EXMPLE: TOUX.FIEVRE.CRACHA.RALES/HILEOP.TRAMEA.FIEVRE*
 .A.-B.C/A.-B.-D/-B.C.-D*
 NOUVEAU TRT(=0) OU FIN? ..0

F. NOM DU SYST.? =AXIOME
 CREATION=1,MISE A JOUR=2,MISE A.J.AVEC RAZ=3,ANNULATION=4 .2
 QUE FAIRE SUR CE GROUPE
 CREATION=1,MISE A JOUR=2,MISE A.J.AVEC RAZ=3,ANNULATION=4 .1
 NOM DU SYSTEME D AXIOMES A CREER:DROITE
 SYNTAXE D'ECRITURE DES AXIOMES-C'EST A DIRE DES AFFIRMAIONS DE BASES
 LES SIGNES(OU ATOMES)SONT IDENTIFIES PAR 6 CARACT. AU PLUS,LE 1E EST TOUJOU
 LES AXIOMES SONT DES SUITES DE SIGNES,OU DE SIGNES NIES (PRECEDES DU SYMBOLE
 LES AXIOMES SONT SEPARES PAR LE CARACTERE /
 LE DERNIER AXIOME EST TERMINE PAR LE CAR. *
 EXMPLE: TOUX.FIEVRE.CRACHA.RALES/HILEOP.TRAMEA.FIEVRE*
 .-A.B.-D/-A.B.-C*
 NOUVEAU TRT(=0) OU FIN? ..FI

2.3.2. DEMONS

Le programme de démonstration travaille en deux temps.

2.3.2.1. Prise en considération des signes du patient.

2.3.2.2. Démonstration par la méthode de DAVIS et PUTNAM.

2.3.2.1. Prise en considération des signes du patient

Cette partie du programme commence par ouvrir le fichier CATDEM où sont répertoriés tous les noms des groupes des systèmes d'axiomes.

Après cette ouverture de fichier, le programme demande sur la visu le nom du groupe d'axiomes c'est-à-dire le nom de la pathologie à traiter.

Le programme appelle ensuite RGPAX1 afin de vérifier que le nom de groupe existe réellement et afin de se positionner dans CATDEM sur ce groupe. Si le nom ne se trouve pas dans CARDEM, on envoie un message.

DEMONS demande alors si on veut travailler sur un groupe d'axiomes bien déterminé (SWITCH =1) ou si on opte pour le travail sur tous les systèmes (SWITCH =2). En effet, on peut confronter les signes du patient avec un diagnostic ou avec tous les diagnostics relatifs à une pathologie.

Si le SWITCH =(1), la visu affiche une demande du numéro de système à traiter.

Si l'utilisateur répond 0, c'est qu'il désire le défilement des noms de fichier.

Il choisira le numéro séparé par après.

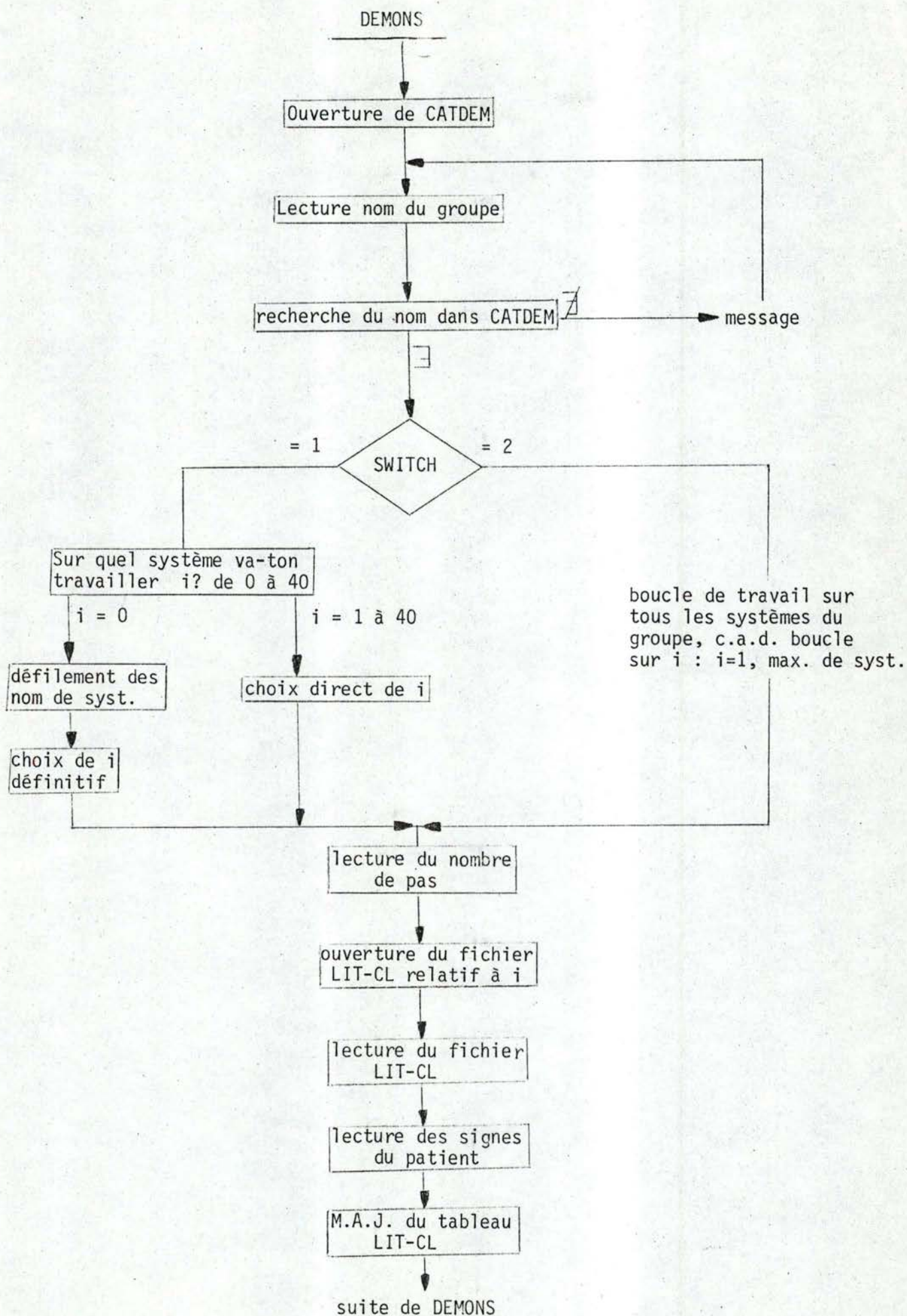
Si le SWITCH =(2), la démonstration sera faite par rapport à tous les fichiers représentant les systèmes d'axiomes.

On demande ensuite le nombre de pas nécessaires à la démonstration.

Ce nombre de pas servira, à la création du fichier FSPLIT, à l'éclatement du système d'axiomes. Nous verrons l'utilité de ce fichier lors de l'explication de la démonstration proprement dite.

DEMONS ouvre ensuite le fichier LIT-CL choisi (1) ou un de tous les fichiers LIT-CL d'une pathologie (2) (l'un à la suite de l'autre) et effectue une lecture préalable de ce fichier.

C'est ici que l'utilisateur introduit les signes du patient par la sous-routine INDT. Ces signes seront introduits dans les tableaux LIT-CL relatifs au diagnostic testé de la même manière que pour le programme ENTRAX avec les deux paramètres IN et KL égaux à 2.



Il est à remarquer que l'on ne modifie pas les fichiers qui sont sur disque car ils constituent la connaissance médicale. La démonstration se fera donc sur les tableaux LIT-CL relatifs à un diagnostic dans lesquels de nouveaux littéraux et clauses ont été ajoutés.

2.3.2.2. Démonstration par la méthode de DAVIS et PUTNAM

DEMONS a déjà, dans un premier temps, ajouté les signes du patient aux tableaux LIT-CL traités. Par la méthode de DAVIS et PUTNAM, nous allons essayer de prouver que l'ensemble formé par les clauses de la connaissance médicale (obtenue par négation) et les clauses formées par les signes du patient est inconsistant. Si nous arrivons à en prouver l'inconsistance, c'est que les signes du patient entraînent effectivement le diagnostic relatif à la connaissance médicale du fichier LIT-CL traité.

Nous allons nous attarder maintenant sur la manière dont cette méthode a été mise en oeuvre dans DEMONS.

Le travail va s'effectuer sur une table LIT-CL complétée et avec un fichier de travail FSPLIT.

Le programme commence par ouvrir ce fichier de travail et calcule le nombre de colonnes maximum dans les deux tableaux LIT-CL (économie dans les boucles de travail sur les tableaux).

Il applique ensuite les 3 premières règles de DAVIS et PUTNAM, c'est-à-dire :

- a. Tautologie
- b. Littéral pur
- c. Littéral unique.

Ces trois règles sont appelées par la sous-routine DAVIS.

a. Tautologie (sous-routine Tautol)

Cette recherche se fait dans la table LIT. En effet, si sur une même ligne (c'est-à-dire un même littéral), un même numéro de clause apparaît au moins une fois en positif et une fois en négatif, c'est que cette clause est une tautologie.

Rappel : une clause est une disjonction finie de zéro ou plusieurs littéraux

Le fait d'avoir un même élément nié et non nié dans une disjonction de littéraux entraîne que la clause est toujours valide (tautologie).

Si une tautologie est ainsi découverte, DEMONS appelle EFCLAU qui est une sous-routine d'effacement de la clause :

- efface la clause de la table CL,
- efface le numéro de cette clause partout où elle intervient dans la table LIT.

b. Littéral pur (sous-routine RLP)

RLP appelle d'abord PURLIT qui est une sous-routine de recherche des littéraux purs c'est-à-dire les littéraux apparaissant toujours sous le même signe dans les clauses.

PURLIT prend les littéraux un à un dans une boucle sur toutes les lignes non vierges de LIT. Il prend le premier numéro de clause et en regarde le signe. Il prend ensuite les numéros suivants toujours relatifs au même littéral et compare leur signe par rapport au premier.

- Si le signe est identique, on continue la recherche avec le même littéral pour les numéros de clause suivants et ainsi de suite jusqu'au moment où le numéro de clause devient nul. On a alors trouvé un littéral pur.
- Si le signe est différent, on passe au littéral suivant.

PURLIT remet en paramètre à RLP le numéro de littéral pur s'il en existe un et met zéro en paramètre s'il n'en n'existe pas.

Si le paramètre est différent de zéro, alors RLP fait appel à EFCLAU pour effacer toutes les clauses contenant ce littéral pur et on supprime ce littéral pur de la table LIT.

- Si l'ensemble des clauses est vide après ces suppressions, c'est que le système est satisfaisable (IAX = 2).
- Sinon, on continue la démonstration.

Remarque : IAX est le paramètre transmis au programme principal en rendant compte du résultat des opérations effectuées quant à la consistance ou l'inconsistance du système.

c. Littéral unique (sous-routine UNILIT)

Cette sous-routine recherche les clauses monolittérales. Ce travail est rendu aisé car la première colonne de CL contient le nombre de littéraux compris dans la clause.

UNILIT parcourt donc la table CL en consultant cette première colonne jusqu'au moment où celle-ci est égale à 1.

On efface la ligne relative à ce littéral dans CL (EFCLAU) après avoir sauvé son numéro. On recherche alors dans toutes les autres clauses ce littéral sauvé.

- Si on trouve un littéral identique, deux situations peuvent se présenter :
 - Ce littéral est de même signe, on peut alors effacer cette clause (EFCLAU)
 - Ce littéral est de signe opposé, on s'occupe de ce littéral complémentaire en le supprimant de toutes les clauses le contenant. Pour cela, la sous-routine fait appel à la sous-routine EFLIT.

Si cette suppression rend vide un ensemble de clauses, c'est que le système est inconsistant. En effet, on se trouve alors devant une expression du type $(D \cdot \sim D)$. On fait parvenir cette information au programme central par le paramètre IAX.
- Si on ne trouve pas de clause contenant le littéral sauvé, on revient au programme principal pour l'application de la 4ème règle.

Ceci termine la sous-routine DAVIS qui mettait en application les trois premières règles de DAVIS et PUTNAM.

Nous allons maintenant appliquer la 4ème règle de DAVIS et PUTNAM c'est-à-dire la règle du PARTAGE.

Pour pouvoir appliquer cette règle, le littéral doit apparaître au moins 2 fois : normal et complémenté.

Le programme va donc balayer le tableau des littéraux en bouclant sur toutes les lignes non nulles de la table LIT. Pour un littéral, il prend le premier numéro de clauses où il intervient et compare son signe avec les signes des numéros de clause suivants.

- S'ils sont tous de même signe, on passe au littéral suivant. Si on ne peut appliquer la règle 4, on envoie un message "pas de solution".
- Sinon, on peut appliquer la 4ème règle de DAVIS et PUTNAM comme suit.

C'est ici qu'intervient le fichier de travail FSPLIT que nous avons créé et ouvert au début du programme. La dimension de ce fichier a été définie par la lecture du nombre de pas désiré (voir § 2.3.2.1.).

Voyons maintenant l'utilité de ce fichier de travail FSPLIT. Ce fichier sert à dédoubler les tables LIT-CL (Partage) en tables TL1-TC1 et TL2-TC2 de la façon suivante.

Le programme recopie d'abord les tables LIT-CL dans les tables TL1-TC1 dans lesquelles il va effacer le littéral choisi apparaissant sous forme normale et complémentée (appel à la sous-routine EFLIT).

Lorsque ceci a été fait, il recopie les tableaux TL1-TC1 obtenus dans les tables TL2-TC2 appartenant aussi à FSPLIT.

DEMONS va ensuite différencier les tables TL1-TC1 et TL2-TC2 en séparant les ensembles de clauses comme la règle 4 nous l'indique. Il regarde alors dans la table LIT-CL quand le littéral traité apparaît positif et négatif dans les clauses du système (recherche dans LIT)

- quand il apparaît positif dans une clause de LIT-CL, on supprime toutes les clauses où le littéral apparaissait sous forme négative.

Pour cela, on fait appel à SUPCL sur les tables TL1-TC1. (1)

- quand il apparaît négatif dans une clause de LIT-CL, on supprime toutes les clauses où le littéral apparaissait sous forme positive. Pour cela, on fait appel à SUPCL sur les tables TL2-TC2. (2)

Exemple de partage

Soit $S = (S+R) \cdot (\sim S+T) \cdot (S+U+R) \cdot (\sim S+\sim W) \cdot X$.

Mettons ceci sous forme de tableaux LIT-CL.

Table des littéraux (LIT)

* R	* 2	* 1	3			
* S	* 4	* 1	-2	+3	-4	
* T	* 1	* 2				
* U	* 1	* 3				
* W	* 1	* -4				
* X	* 1	* 5				

Table des clauses (CL)

```

* 2 * 2 1
* 2 * -2 3
* 3 * 2 4 1
* 2 * -2 -5
* 1 * 6

```

Nous obtenons ensuite la première génération des tables de travail en supprimant le littéral S nié et non nié des tableaux de départ LIT-CL. LIT-CL demeure inchangé.

Tables TL1-TC1 et TL2-TC2- TL1 et TL2

```

* R * 2 * 1 3
* S * 0 *
* T * 1 * 2
* U * 1 * 3
* W * 1 * -4
* X * 1 * 5

```

- TC1 et TC2

```

* 1 * 1
* 1 * 3
* 2 * 4 1
* 1 * -5
* 1 * 6

```

Après application du point (1), nous allons obtenir les tables TL1-TC1 suivantes.

La table TL1 reste inchangée mais la table TC1 devient la suivante :

TC1

```

* 1 * 1
* 2 * 4 1
* 1 * 6

```

d'où l'ensemble de clauses de la première partie
du partage est $\{R, U + R, X\}$ ou encore $R.(U+R).X (S')$

Après application du point (2), nous obtenons les tables TL2-TC2 suivantes.

La table TL2 reste inchangée mais la table TC2 devient la suivante :

TC2

* 1 * 3 d'où l'ensemble de clauses de la deuxième partie du partage
 * 1 * -5 est : $\{T, \sim W, X\}$ ou encore $T \cdot \sim W \cdot X$ (S")
 * 1 * 6

L'ensemble de départ S a donc été séparé en deux sous-ensembles S' et S" dont nous devons prouver l'inconsistance.

Si un seul de ces sous-ensembles est consistant, l'ensemble S de départ l'est également.

Le programme écrit ensuite ces nouvelles tables, sur disque, dans le fichier FSPLIT et se sert d'un vecteur (MARK) pour mémoriser les tables ainsi créées et pour savoir si elles ont déjà été traitées ou non. La dimension de MARK est également relative au nombre de pas demandé. Un élément de ce tableau, MARK (i) peut prendre plusieurs valeurs :

- = 1 si on a créé les tables i et pas encore traitées
- = -1 si on a créé et déjà traité la table correspondante
- 0 si pas encore créée.

La première table du fichier FSPLIT est alors lue et on lui applique de nouveau les 3 premières règles de DAVIS et PUTNAM après avoir positionné le vecteur MARK.

Après les avoir appliquées, le programme est switché pour ne plus recréer directement de nouvelles tables de travail. C'est à ce moment que le programme teste un paramètre IAX. Ce dernier est généré comme on l'a vu à la sortie des différentes sous-routines et relate la consistance ou l'inconsistance du système traité.

Ainsi IAX = 0 : pas encore de conclusions

IAX = 1 : système inconsistant

IAX = 2 : ensemble est satisfaisable après effacement des clauses.

- Si IAX = 0, alors d'autres tables de travail TL1-TC1 et TL2-TC2 sont générées pour réappliquer les 3 premières règles de DAVIS et PUTNAM. On applique un pas de plus à la démonstration si le nombre maximum n'a pas été atteint.

- Si $IAX = 1$, c'est que le sous-ensemble traité est inconsistant. Si l'autre sous-ensemble n'a pas encore été traité, il faut réappliquer les règles 1,2 et 3 sur l'autre sous-ensemble.

Si les 2 sous-ensembles ont déjà été traités, il est nécessaire de remonter dans les tables afin de les analyser. Ceci peut se faire grâce à la consultation du vecteur MARK.

Ceci se comprend plus aisément si on représente la génération des sous-ensembles de clauses par une arborescence.

- Soit S : un ensemble de clauses représentés sur LIT-CL (1)
- Par application de la règle 4, nous obtenons 2 sous-ensembles S_1 (représenté par TL1-TC1) et S_2 (représenté par TL2-TC2) (2).
- On travaille d'abord sur S_1 pour en montrer l'inconsistance. On ira donc lire sur disque les tables TL1-TC1 correspondantes. Après analyse, supposons qu'on ne puisse pas encore se prononcer sur son inconsistance et que l'on puisse réappliquer la règle 4. Ceci va donc nous donner deux nouveaux sous-ensembles S_{11} et S_{12} .

S_{11} étant représenté par un nouveau TL1-TC1 et S_{12} par TL2-TC2 (4).

On essaie alors de démontrer l'inconsistance de S_{11} . On va alors aller rechercher sur disque le sous-ensemble correspondant au TL1-TC1 traité. On le lit et on l'analyse par les 3 premières règles. Supposons que l'on puisse en démontrer l'inconsistance. On recherche alors la table correspondante de même niveau. On la lit et on l'analyse suivant le même processus.

Supposons de nouveau ce sous-ensemble inconsistant (5).

- C'est à ce moment qu'il est nécessaire de remonter dans les tables pour parcourir l'autre branche de l'arborescence car tous les sous-ensembles doivent être inconsistants pour démontrer l'inconsistance de l'ensemble de départ. Ceci se fait par la consultation de MARK. On analysera alors la 2ème branche principale de l'arborescence comme la première.
(voir schéma p. 136).

- Si $IAX = 2$, c'est-à-dire qu'un sous-ensemble est consistant d'où l'ensemble global des clauses ne peut être inconsistant. On retourne alors au programme principal sans analyser les autres branches de l'arborescence.

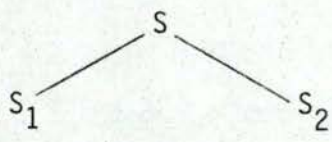
Il se peut aussi que le nombre de pas prévu soit insuffisant pour obtenir un résultat. Un message est alors affiché.

S

LIT-CL

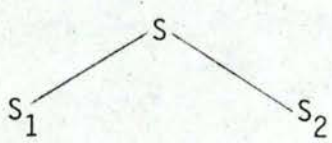
FSPLIT

MARK



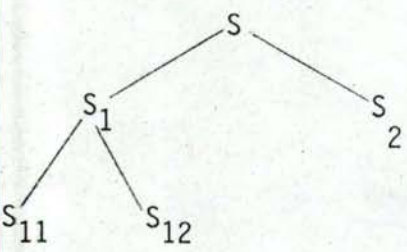
TL1 - TC1
TL2 - TC2

1
1
0
⋮



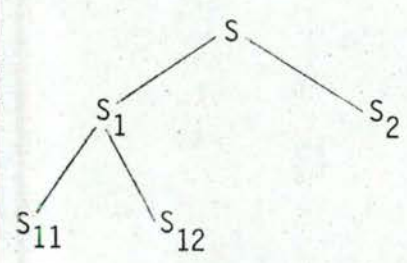
TL1 - TC1
TL2 - TC2

-1
1
0
⋮



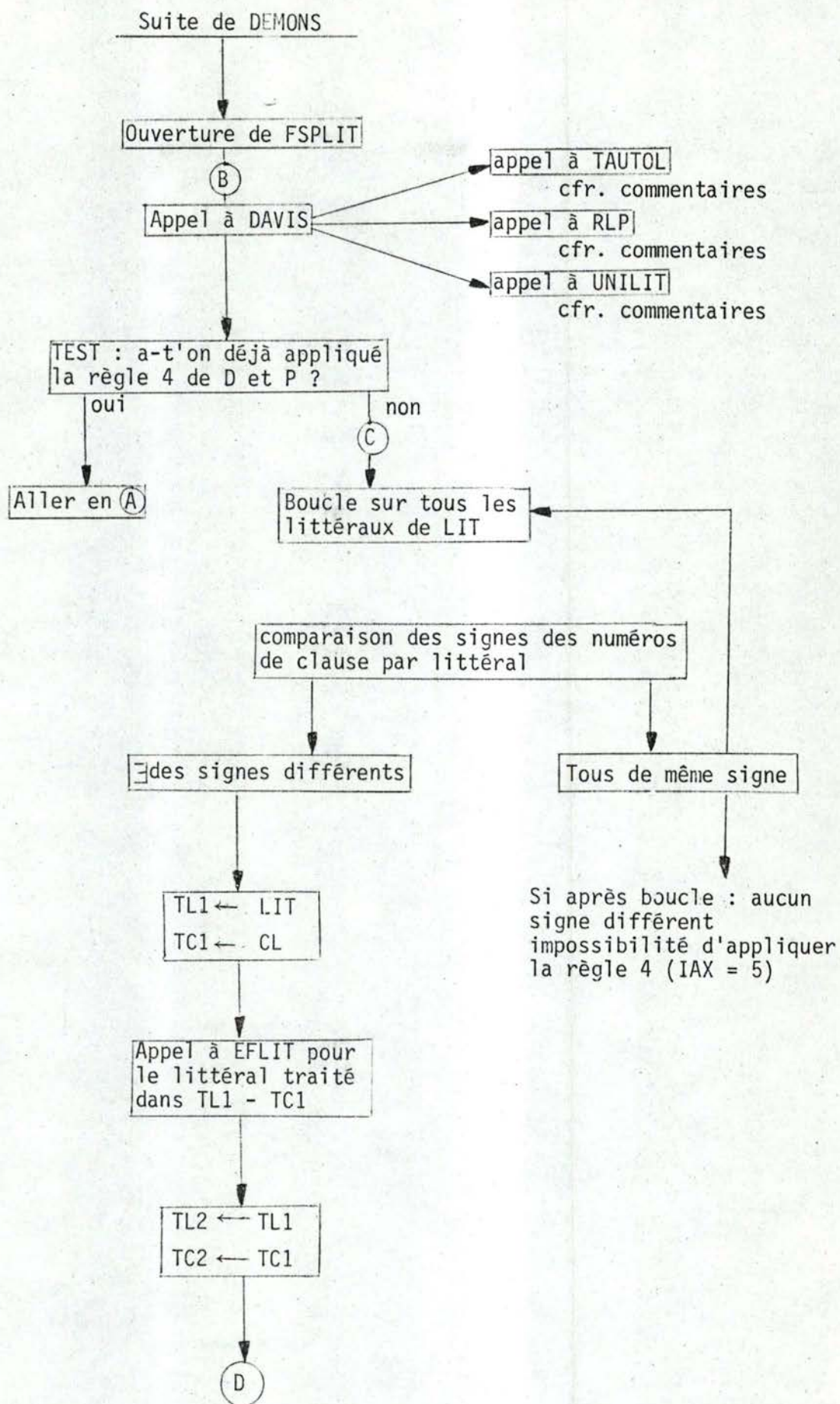
TL1 - TC1
TL2 - TC2
TL1 - TC1
TL2 - TC2

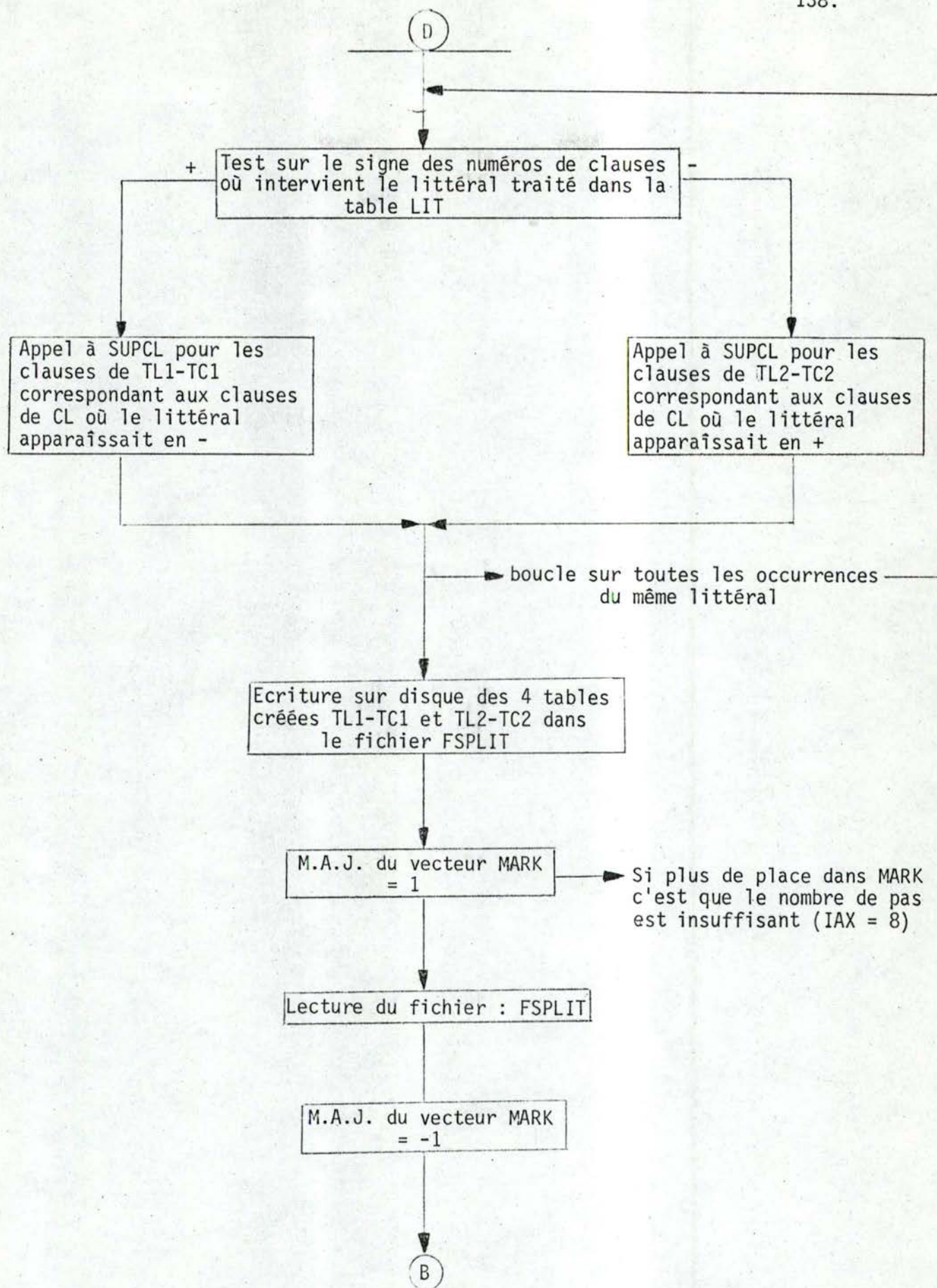
-1
1
1
1
0
⋮

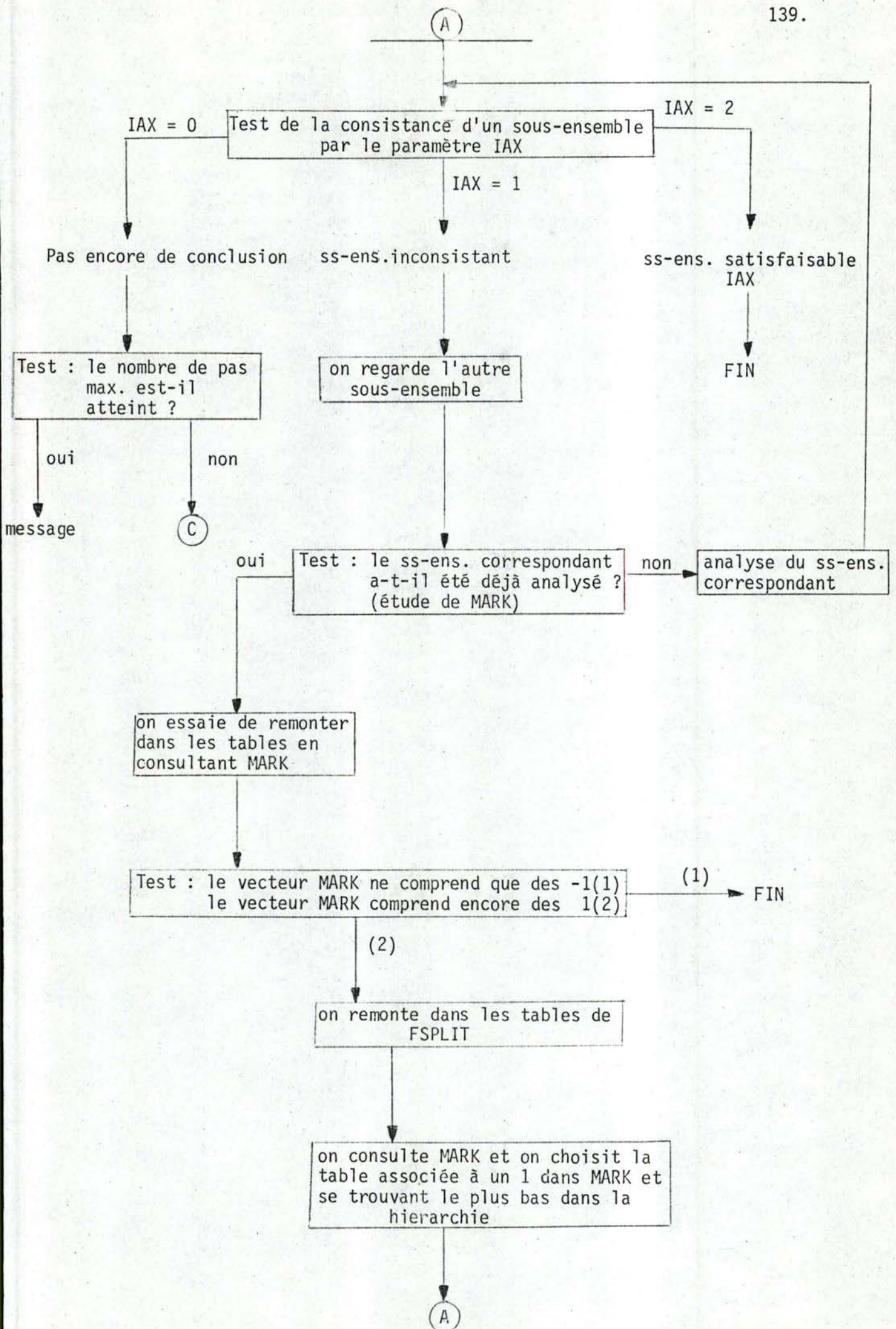


TL1 - TC1
TL2 - TC2
TL1 - TC1
TL2 - TC2

-1
1
-1
-1
0
⋮







2.3.2.3. Amélioration possible du programme de démonstration

L'aide au diagnostic par la méthode de démonstration automatique de théorèmes offre les mêmes résultats que la méthode booléenne étant donné que les mêmes signes et les mêmes fonctions sont choisis dans les deux méthodes.

La méthode booléenne a été choisie au CHUTimone pour sa plus grande vitesse de réponse. En effet, elle ne demande pas, comme la méthode de démonstration automatique, la génération de nombreux tableaux intermédiaires, ce qui requiert une plus grande place mémoire et ce qui occasionne évidemment une perte de temps.

Néanmoins, la méthode de démonstration automatique de théorèmes peut s'avérer intéressante par l'amélioration que nous pouvons y apporter.

Cette amélioration offre l'intérêt en cas de symptomatologie non caractéristique, de faire proposer par l'ordinateur la recherche ou la vérification de certains signes qui permettraient de conclure.

Cette amélioration s'intègre facilement dans le programme DEMONS qui utilise la méthode de DAVIS et PUTNAM.

Afin de mieux visionner la démarche utilisée, nous allons reprendre l'illustration du cas simple de diagnostic vue antérieurement.

Nous rappelons que nous avons deux groupes diagnostiques possibles D_1 et D_2 qui avaient respectivement pour ensemble de clauses

$\{(\neg a, b, \neg c), (a, b, d), (b, \neg c, d)\}$
et $\{(a, \neg b, d), (a, \neg b, c)\}$.

Nous opérons toujours comme nous dicte la méthode de démonstration en essayant de démontrer l'inconsistance de l'ensemble formé par les signes du patient et la connaissance médicale niée relative au diagnostic choisi (sous forme axiomatisée).

La démarche utilisée permettant de mettre à jour les signes non relevés ou non caractéristiques par rapport au diagnostic choisi repose sur un ordre dans l'étude des signes présentés par le patient et à posteriori dans leur recherche ou leur vérification.

Définissons cet ordre.

Les symptômes caractéristiques du diagnostic étudié sont les premiers traités. Ce sont les signes qui se retrouvent toujours niés dans les clauses de la théorie. On leur applique la règle II de DAVIS et PUTNAM, ce qui les fait disparaître de l'ensemble des clauses. On choisit cet ordre car on essaie de garder dans l'ensemble des clauses les signes non caractéristiques du diagnostic étudié.

On poursuit alors la démonstration sur les signes restants en choisissant ceux qui sont le plus souvent niés dans l'ensemble restant. Ce sont les signes qui, souvent, caractérisent le diagnostic étudié. Si l'on effectue une vérification des signes, ces derniers ne seront testés qu'à la fin car ils ne sont pas toujours en contradiction avec le diagnostic étudié. Pour une éventuelle vérification, ces signes sont mémorisés.

Au fur et à mesure de la démonstration, les signes restants sont de moins en moins représentatifs du diagnostic étudié et sont mémorisés dans cet ordre pour être vérifiés les premiers.

Cette démarche permet également de mettre à jour les signes non relevés. En effet, après avoir travaillé sur les signes du patient les moins caractéristiques, d'autres signes peuvent encore subsister dans les ensembles de clauses. Ce sont les signes non relevés dont on devra tenir compte avant même d'opérer les vérifications sur les autres signes.

Remarque : tous les signes non relevés par rapport au diagnostic choisi ne se retrouvent peut-être pas dans l'ensemble des clauses restant après l'application de la règle II de DAVIS et PUTNAM sur les littéraux uniques que sont les signes du patient. En effet, certains de ces signes ont pu être éliminés lors de l'effacement d'une clause.

A la fin de la démonstration, deux cas se présentent :

- Le système est inconsistant, aucune vérification n'est à effectuer car la symptomatologie est bien caractéristique du diagnostic testé.
- Le système est satisfaisable, d'où la symptomatologie n'est pas caractéristique du diagnostic testé. Or on croit que le patient présente réellement le diagnostic étudié et on désire vérifier les signes non caractéristiques. (Une erreur a pu être commise lors du remplissage du questionnaire ou à l'introduction des signes en machine).

On prendra d'abord en considération les signes non relevés avant de vérifier les signes présentés par le patient dans l'ordre inverse de leur mémorisation.

Prenons deux exemples de symptomatologie non caractéristiques du diagnostic D_1 reflétant la démarche expliquée ci-avant :

- Soit la symptomatologie du patient A : b et $\sim c$
- Soit la symptomatologie du patient B : $a, b, \sim c, \sim d$.

Etudions d'abord le cas du patient A.

Formons d'abord l'ensemble composé des signes du patient et de la connaissance médicale relative au diagnostic D_1 , ce qui donne :

$$S = \{b, \sim c, (a + \sim b + c), (\sim a + \sim b + \sim d), (\sim b + c + \sim d)\}.$$

Les signes 'b' et ' $\sim c$ ' sont caractéristiques du diagnostic D_1 car on les retrouve niés dans les clauses de la connaissance médicale.

On leur applique la règle II de DAVIS et PUTNAM ce qui donne :

$$S_1 = \{a, \sim a + \sim d, \sim d\}.$$

Après avoir travaillé sur les signes des patients, les signes a et d subsistent dans l'ensemble S réduit. Ces signes sont à rechercher afin de porter un diagnostic valable sur le patient A. Il est néanmoins à remarquer qu'un seul des signes soit ' $\sim a$ ' soit 'd' suffit pour aboutir à la conclusion d'inconsistance du système.

Etudions maintenant le cas du patient B.

Formons d'abord l'ensemble de départ :

$$S = \{a, b, \sim c, \sim d, (a + \sim b + c), (\sim a + \sim b + \sim d), (\sim b + c + \sim d)\}.$$

Les signes 'b' et 'c' étant caractéristiques du diagnostic D_1 (toujours niés dans les ensembles de clauses), nous leur appliquons la règle II (règle du littéral unique) de DAVIS et PUTNAM. Nous obtenons ainsi :

$$S_1 = \{a, \sim d, a, \sim a + \sim d, \sim d\}.$$

Il reste le signe 'a' qui est nié une fois sur deux dans les clauses du diagnostic D_1 et le signe 'd' qui se retrouve sous le même signe que dans les clauses de D_1 .

Le signe 'd' n'est donc jamais caractéristique du diagnostic D_1 , c'est pourquoi nous analysons d'abord le signe a par la règle II après l'avoir mémorisé.

$$S_2 = \{\sim d, \sim d, \sim d\}.$$

L'ensemble de départ est donc satisfaisable. Le signe 'd' est également mémorisé.

On peut alors afficher les signes du patient non caractéristiques par rapport au diagnostic étudié. Le premier signe à vérifier est ainsi affiché, il s'agit de 'd'. Si la vérification entraîne une modification du signe, la nouvelle symptomatologie du patient est testée.

Sinon, on remonte dans la vérification. On en arrive ainsi au signe 'a' où l'on effectue la même vérification.

Si le signe a été modifié, on teste à nouveau la symptomatologie du patient, sinon on s'en tient à notre ancien diagnostic "symptomatologie du diagnostic D_1 non caractéristique".

Cette amélioration s'avère utile lorsque le diagnostic "automatique" ne rencontre pas celui proposé par le médecin. Or ce dernier étant quasi sûr de son choix tiendra à savoir quels sont les signes manquants ou contradictoires par rapport à son diagnostic.

Nous verrons des exemples relatifs à la pathologie thyroïdienne illustrant mieux encore cette amélioration de la méthode.

2.3.2.4. Applications du programme de démonstration : DEMONS

Nous avons vu lors de l'étude de ENTRAX, l'élaboration des ensembles de clauses représentant la connaissance médicale.

Nous les avons mémorisés dans des fichiers portant des noms significatifs par rapport au diagnostic étudié.

Le but de DEMONS est d'entrer en machine les signes du patient et de les diagnostiquer par rapport aux ensembles de clauses créés par ENTRAX.

Le programme DEMONS livre le résultat de la démonstration par un paramètre (IAX).

Suivant ce paramètre, nous sommes capables de déterminer la consistance ou l'inconsistance de l'ensemble de clauses obtenu en adjoignant les signes du patient à la théorie médicale relative à un diagnostic.

- Si nous arrivons à la solution d'inconsistance du système étudié, c'est que la symptomatologie du patient est caractéristique du diagnostic étudié. Nous affichons alors sur la visu le commentaire suivant :

"LES SIGNES PRESENTES PAR LE PATIENT ENTRAINENT LE DIAGNOSTIC DE ..."

où les "..." représentent le nom du fichier contenant l'ensemble de clauses relatif au diagnostic étudié.

- Si nous arrivons à la solution de consistance du système étudié, c'est que la symptomatologie du patient n'est pas caractéristique du système étudié.

Nous affichons alors sur la visu le commentaire suivant :

"LES SIGNES PRESENTES PAR LE PATIENT N'ENTRAINENT PAS LE DIAGNOSTIC DE ..."

où les "..." représentent le nom du fichier contenant l'ensemble de clauses relatif au diagnostic étudié.

a. Application simple sur le groupe AXIOME créé par ENTRAX

Nous rappelons que le groupe AXIOME comprend deux ensembles de clauses représentatifs de la connaissance médicale relative à une pathologie imaginaire. GAUCHE et DROITE sont les noms des deux diagnostics existants et des fichiers les supportant.

Voyons maintenant quelques exemples simples montrant les possibilités de dialogue et de résultat qu'offrent ce programme de démonstration.

Exemple A du listing ci-joint

On désire tester la symptomatologie du patient par rapport aux deux systèmes d'axiomes existants. On répond donc (2) à la question :

"TRAVAIL SUR UN SEUL SYSTEME (1) OU SUR TOUS LES SYSTEMES DU GROUPE (2)" :

Le nom du système traité s'affiche et il faut alors taper sur le clavier la symptomatologie du patient où le "." représente le "ou", le "/" représente le "et", le "-" représente la négation et le "*" signifie que le dernier signe du patient a été pris en compte.

Dans ce premier exemple, on teste si le signe '-A' suffit pour diagnostiquer le patient. Cette symptomatologie rudimentaire ne parvient à caractériser aucun des deux diagnostics. C'est ainsi que s'affichent les commentaires suivants :

"LES SIGNES PRESENTES PAR LE PATIENT N'ENTRAINENT PAS LE DIAGNOSTIC DE GAUCHE"
(idem pour le diagnostic de DROITE).

Exemple B

On teste une autre symptomatologie de patient par rapport aux deux diagnostics possibles.

On tape au clavier : -A/B/-C* c'est-à-dire que trois signes ont été pris en compte pour ce malade : les signes -a, b et -c.

Or nous avons vu lors de l'élaboration des systèmes d'axiomes du groupe AXIOME que l'une des associations caractéristiques du diagnostic GAUCHE était :

$$\sim a \cdot b \cdot \sim c \longrightarrow \text{GAUCHE } (D_1).$$

La symptomatologie entrée en machine dans ce point B devrait donc entraîner le diagnostic GAUCHE.

C'est ce que confirme le commentaire apparaissant à l'écran.


```

TRAVAIL DEMANDE .....DEMONS
NOM DU SYSTEME D AXIOMES A TRAITER .....AXIOME
NBE DE PAS ..5
TRAVAIL SUR 1 SEUL SYSTEME(1) OU SUR TOUS LES SYSTEMES DU GROUPE(2) : .2
SYSTEME TRAITE: GAUCHE
NEGATION A DEMONTRER
.-A*
LES SIGNES PRESENTES PAR LE PATIENT N ENTRAINENT PAS LE DIAGNOSTIC DE GAUCHE
SYSTEME TRAITE: DROITE
NEGATION A DEMONTRER
.-A*
LES SIGNES PRESENTES PAR LE PATIENT N ENTRAINENT PAS LE DIAGNOSTIC DE DROITE
NOUVEAU TRT(=0) OU FIN? ..0
NOM DU SYSTEME D AXIOMES A TRAITER .....AXIOME
NBE DE PAS ..5
TRAVAIL SUR 1 SEUL SYSTEME(1) OU SUR TOUS LES SYSTEMES DU GROUPE(2) : .2
SYSTEME TRAITE: GAUCHE
NEGATION A DEMONTRER
.-A/B/-C*
LES SIGNES PRESENTES PAR LE PATIENT ENTRAINENT LE DIAGNOSTIC DE GAUCHE
SYSTEME TRAITE: DROITE
NEGATION A DEMONTRER
.-A/B/-C*
LES SIGNES PRESENTES PAR LE PATIENT N ENTRAINENT PAS LE DIAGNOSTIC DE DROITE
NOUVEAU TRT(=0) OU FIN? ..0
NOM DU SYSTEME D AXIOMES A TRAITER .....AXIOME
NBE DE PAS ..5
TRAVAIL SUR 1 SEUL SYSTEME(1) OU SUR TOUS LES SYSTEMES DU GROUPE(2) : .1
NO DU GROUPE _ 0:DEFILEMENT _ 1.40:POSITIONNEMENT ..0
1 GAUCHE
2 DROITE
NUM.CHOISI?= ..1
SYST.D AXIOMES CHOISIS GAUCHE
NEGATION A DEMONTRER
.-A/B/-C/D*
LES SIGNES PRESENTES PAR LE PATIENT ENTRAINENT LE DIAGNOSTIC DE GAUCHE
NOUVEAU TRT(=0) OU FIN? ..0
NOM DU SYSTEME D AXIOMES A TRAITER .....AXIOME
NBE DE PAS ..5
TRAVAIL SUR 1 SEUL SYSTEME(1) OU SUR TOUS LES SYSTEMES DU GROUPE(2) : .1
NO DU GROUPE _ 0:DEFILEMENT _ 1.40:POSITIONNEMENT ..1
SYST.D AXIOMES CHOISIS GAUCHE
NEGATION A DEMONTRER
.-A/B/-C/-D*
LES SIGNES PRESENTES PAR LE PATIENT ENTRAINENT LE DIAGNOSTIC DE GAUCHE
NOUVEAU TRT(=0) OU FIN? ..0
NOM DU SYSTEME D AXIOMES A TRAITER .....AXIOME
NBE DE PAS ..5
TRAVAIL SUR 1 SEUL SYSTEME(1) OU SUR TOUS LES SYSTEMES DU GROUPE(2) : .2
SYSTEME TRAITE: GAUCHE
NEGATION A DEMONTRER
.A/-B/D*
LES SIGNES PRESENTES PAR LE PATIENT N ENTRAINENT PAS LE DIAGNOSTIC DE GAUCHE
SYSTEME TRAITE: DROITE
NEGATION A DEMONTRER
.A/-B/D*
LES SIGNES PRESENTES PAR LE PATIENT ENTRAINENT LE DIAGNOSTIC DE DROITE
NOUVEAU TRT(=0) OU FIN? ..FI

```


Exemple C

Le travail s'effectue ici sur un seul système d'axiomes, ce qui se fait très souvent. En effet, le médecin voulant être confirmé dans son diagnostic, propose la symptomatologie de son patient à l'ordinateur pour le confronter à son propre diagnostic. Il demande le défilement des noms de fichiers pour choisir le numéro correspondant au système à tester.

Le résultat proposé confirmera son diagnostic ou le remettra en question. S'il est remis en question, l'amélioration proposée au point précédent sera très utile. En effet, elle permettra d'attirer l'attention du médecin sur des signes absents ou contradictoires par rapport à son diagnostic proposé.

Les cas D et E sont d'autres exemples similaires.

b. Application sur le groupe THYROIDE créé par ENTRAX

Nous avons vu lors de l'étude du programme de démonstration DEMONS, que celui-ci travaillait en deux temps. En effet, dans une première partie, DEMONS prend en considération les signes d'un patient et dans la seconde, opère la démonstration par rapport à un diagnostic choisi.

Revenons d'abord sur la première partie de DEMONS qui prête à la discussion. Le programme tel qu'il a été conçu prévoit une introduction manuelle des signes du patient, c'est-à-dire qu'une opératrice frappe sur le clavier les quelques signes retenus.

Or, cette démarche pourrait se faire automatiquement à partir du dossier médical d'un patient créé sur base d'un questionnaire (voir en annexe 1 quelques pages d'exemple de ce formulaire) et sur base d'examens paracliniques opérés sur le patient. Les quelques signes retenus par nos systèmes d'axiomes relatifs à la pathologie thyroïdienne sont compris dans le dossier SYCVAR. Il est possible de les y extraire par programme, ce qui supprimerait l'intervention manuelle d'une opératrice pour proposer ces signes au diagnostic.

Ceci permettrait évidemment un gain de temps énorme tout en supprimant les erreurs de manipulations intermédiaires. Le temps ne nous a malheureusement pas permis de terminer cet échange d'informations entre le dossier SYCVAR et le programme DEMONS. Il est à noter qu'un programme similaire existe entre le dossier SYCVAR et le programme de diagnostic booléen où les signes repris dans les fonctions booléennes sont sélectionnés et analysés.

Nous nous sommes donc contentés de tester DEMONS en reprenant certains résumés de dossiers (voir exemple en annexe 2) avec le diagnostic correspondant obtenu par la méthode de logique booléenne.

Nous allons montrer maintenant quelques exemples de diagnostic. Prenons l'exemple marqué A sur le listing ci-joint. Dans cet exemple, on désire tester deux symptomatologies différentes par rapport au diagnostic d'hypothyroïde. Nous devons donc sélectionner le fichier relatif au diagnostic choisi : HYP0. La première symptomatologie notée A_1 est composée des signes suivants :

Frilosité	: OUI	(S_1)
Peau sèche et squameuse	: OUI	(S_3)
Taux de fixation 6ème heure diminué	: OUI	(S_5)
Test Blanquet - Meyniel III diminué	: OUI	(S_7)
Voix rauque et nasonnée	: OUI	(S_9)
Taux de fixation 24ème heure augmenté	: NON	($-S_{11}$)
Tachycardie	: NON	($-S_{12}$)
Taux de fixation 42ème heure augmenté	: OUI	(S_{17})

Les signes sont introduits au terminal après avoir été transformés en S numérotés par consultation de la table de correspondance des signes. Après le signe de fin d'écriture "x", la démonstration s'exécute par la création de tables successives LIT et CL et le commentaire relatif au résultat apparaît alors à l'écran. Les signes présentés ici entraînent effectivement le diagnostic d'hypothyroïde.

Voyons maintenant la deuxième symptomatologie (A_2) étudiée dans cet exemple A. Le commentaire relatif à cet exemple nous informe que les signes présentés par le patient n'entraînent pas le diagnostic d'hypothyroïde. Si on analyse le système d'axiomes relatif à l'hypothyroïde, on remarque que c'est par le signe de frilosité absent ($-S_1$) que ce patient n'est pas affecté du diagnostic d'hypothyroïde.

On peut considérer que la démarche prise pour cet exemple A serait celle d'un médecin qui, ayant porté son diagnostic, veut se le faire confirmer par une méthode automatique.

NOM DU SYSTEME D AXIOMES A TRAITERTHYROIDE
 NBE DE PAS ..10
 TRAVAIL SUR 1 SEUL SYSTEME(1) OU SUR TOUS LES SYSTEMES DU GROUPE(2) : .1
 NO DU GROUPE _ 0:DEFILEMENT _ 1.40:POSITIONNEMENT ..0
 1 HYPO
 2 NORMAL
 3 HYPER
 4 THYREO
 NUM.CHOISI?= ..1
 SYST.D AXIOMES CHOISIS HYPO
 NEGATION A DEMONTRER
 .S1/S3/S5/S7/S9/S10/S17/-S11/-S12*
 LES SIGNES PRESENTES PAR LE PATIENT ENTRAINENT LE DIAGNOSTIC DE HYPO
 NOUVEAU TRT(=0) OU FIN? ..0
 NOM DU SYSTEME D AXIOMES A TRAITERTHYROIDE
 NBE DE PAS ..10
 TRAVAIL SUR 1 SEUL SYSTEME(1) OU SUR TOUS LES SYSTEMES DU GROUPE(2) : .1
 NO DU GROUPE _ 0:DEFILEMENT _ 1.40:POSITIONNEMENT ..1
 SYST.D AXIOMES CHOISIS HYPO
 NEGATION A DEMONTRER
 .-S1/S3/S5/S7/S9/-S12*
 LES SIGNES PRESENTES PAR LE PATIENT N ENTRAINENT PAS LE DIAGNOSTIC DE HYF
 NOUVEAU TRT(=0) OU FIN? ..0 0

'autres exemples de diagnostic se trouvent en annexe 3.

Il va donc tester les signes du patient par rapport à son diagnostic.

- Si le résultat confirme son idée, le médecin peut être certain de son choix. En effet, vu la quasi-certitude des implications logiques ayant servi à construire les ensembles de clauses; lorsqu'une des clauses est respectée, le diagnostic associé est quasi-sûr. Autrement dit, si le commentaire de DEMONS est : "LES SIGNES PRESENTES PAR LE PATIENT ENTRAINENT LE DIAGNOSTIC DE ...", ce diagnostic est quasi certain.

C'est le cas de l'exemple A_1 ; un médecin ayant diagnostiqué un patient hypothyroïdien, soumet néanmoins ses signes à la méthode de démonstration automatique en sélectionnant le fichier HYP0. Le commentaire du résultat confirme son opinion, il peut donc porter un diagnostic quasi sûr et certain.

- Par contre, si le résultat contredit le diagnostic du médecin, ce n'est pas pour autant que son diagnostic est faux. En effet, les méthodes de diagnostic automatique sont fragiles; il suffit qu'un seul signe du patient soit contradictoire ou non relevé pour que l'opinion du médecin soit remise en question. L'amélioration proposée plus haut devrait jouer ici un rôle très important car elle nous permet d'obtenir les signes contradictoires et non relevés par rapport au diagnostic étudié.

Il est utile de séparer les signes contradictoires des signes non relevés.

- En effet, s'il existe un ou des signes contradictoires aux symptômes définissant le diagnostic étudié, le choix du médecin est fortement ébranlé. Il sera même amené à revoir son diagnostic en tenant compte de ce signe.

Nous avons un exemple de ce type de problème avec le cas A_2 . Le médecin voulant confirmer son choix s'aperçoit que le signe $-S_1$ (non frilosité) remet son diagnostic en question.

- S'il existe des signes non relevés, le diagnostic automatique est faussé. Le choix de médecin ne peut être remis en question avant d'avoir relevé les symptômes ou examens manquants et d'avoir fait à nouveau tourner le programme de démonstration avec l'ensemble des signes ainsi complété.

Nous pouvons illustrer ce point par l'exemple du dossier 5994 (voir annexe 2).

Nous allons comparer le diagnostic relevé par le dossier 5994 à partir du résumé donné avec notre diagnostic par la méthode de démonstration automatique de théorèmes.

Nous allons tester les signes présents dans le résumé par rapport au fichier HYPO car nous supposons a priori que nous allons aboutir au même diagnostic que par la méthode booléenne.

Parcourons la liste des principaux symptômes et examens paracliniques afin de retenir ceux pris en considération par notre système d'axiomes relatif à l'hypothyroïde.

Ces signes et examens pris en compte sont par ordre de leur numérotation dans la table de correspondance des signes : (ils sont marqués d'une croix dans le listing de l'annexe 2) :

$S_1, S_3, S_4, S_5, S_6, S_9, \sim S_{11}, \sim S_{13}, \sim S_{14}, \sim S_{15}$ et $\sim S_{16}$.

Le commentaire du résultat de DEMONS nous indique que les signes proposés n'entraînent pas le diagnostic d'hypothyroïde.

Lorsque nous appliquons l'amélioration vue plus haut, on peut s'apercevoir que trois signes sont absents pour diagnostiquer : S_2, S_7, S_8 .

En fait, S_7 suffit pour nous amener à diagnostiquer dans le même sens que la méthode booléenne.

On refait alors tourner DEMONS avec les nouveaux signes et on obtient le même résultat que celui proposé par la méthode booléenne.

Ceci renforce le point de vue que l'on avait sur les méthodes de diagnostic automatique. Il s'agit avant tout d'une aide au diagnostic par la recherche des signes litigieux ou absents, et d'une confirmation de l'opinion du médecin. Les différents exemples et explications qui précèdent, montrent que le diagnostic automatique peut difficilement travailler seul et que l'importance du médecin est primordiale.

CONCLUSION

CONCLUSION

Tout au long de ce travail, nous avons vu l'explication et l'utilisation des principales méthodes d'aide au diagnostic. Les résultats acquis sur la pathologie thyroïdienne portent à croire en leur succès.

Néanmoins, nous émettons certaines réserves. Bien que les résultats acquis sur la pathologie thyroïdienne soient brillants, ne les généralisons pas car ils portent sur une application très limitée et bien connue où en général, peu de problèmes se posent.

De plus, les cas diagnostiqués avec certitude sont les plus souvent des cas simples c'est-à-dire ceux que le médecin aurait facilement pu traiter seul. Par contre, lorsque le médecin se trouve confronté à une symptomatologie rare et non caractéristique, la machine rencontre le même problème et n'est généralement pas d'un grand secours. Le patient sera alors orienté vers une branche de rejet ou placé dans un mauvais groupe diagnostique.

L'aide réelle apportée est donc relativement faible. On pourrait néanmoins envisager de débarrasser le médecin du travail de routine sur les cas simples afin de lui permettre de s'attarder sur les symptomatologies complexes. L'ordinateur jouerait donc un rôle d'assistance.

L'aide au diagnostic est sans doute l'application la plus contestée de toutes les applications informatiques en médecine. Par contre, si nous nous orientons vers le dossier médical, les opinions sont favorables. En effet, son utilité ne peut plus être remise en question. L'élaboration même du questionnaire permet une structuration et un approfondissement des connaissances, une standardisation de la terminologie. Son utilisation est devenue indispensable dans les hôpitaux ayant entamé l'informatisation de leur dossier. En plus de son rôle médical, le dossier occupe un rôle prépondérant au point de vue administratif : facturation automatique des actes médicaux, ...

Généralement, le dossier médical n'est donc pas remis en question. Il en est de même pour d'autres applications informatiques telles la surveillance du malade sous monitoring, la gestion hospitalière, ...

Les perspectives d'avenir ne sont pas sombres pour les dossiers médicaux qui sont déjà bien accueillis maintenant. Par contre, le problème reste entier pour l'aide au diagnostic.

Recueillir tel ou tel signe et les associer ne signifie pas diagnostiquer. En effet, le raisonnement médical repose sur la connaissance, mais comporte également une grande part de "feeling" du médecin que l'on ne peut pas toujours inculquer à la machine. De plus, les réticences des médecins au sujet du diagnostic automatisé ne favorisent pas son expansion. Nous croyons qu'un plus grand dialogue médecin-machine permettrait de sortir de cette impasse.

Nous émettons donc certaines réserves quant à l'avenir immédiat du diagnostic assisté par ordinateur.

°
° °

ANNEXES

NOM DU MEDECIN TRAITANT :

EXPLORATION FONCTIONNELLE THYROIDIENNE

N° DOSSIER NODOS

Dans la première case mettre le numéro de code thérapeutique (traitement reçu APRES l'exploration)

CODE THERAPEUTIQUE :	0	Ignorance thérapeutique	5	Traitement iodé (Iode 127)
	1	Irathérapie	6	Médicaments en expérimentation
	2	A.T.S.	7	
	3	Chirurgie	8	
	4	Opothérapie thyroïdienne	9	Divers

NOM DU MALADE NOMFA

NOM D'EPOUSE EPOUS

N° DE SECURITE SOCIALE NOSSO

DATE DE L'EXPLORATION DATE-

ADRESSE : mettre le numéro du département en chiffres (code P.T.T.) puis le nom de la ville en clair et en capitales, enfin le reste de l'adresse complète :
Exemple : 13 - MARSEILLE - 30 Rue de Rome

DATE ET LIEU DE NAISSANCE (la date de naissance en chiffres dans l'ordre : jour,mois,année
exemple : 04 09 1922)

DANAI

VILLE

DEPAR

ANNAI

AGE -

SEXE : le malade est-il du sexe masculin ?

OUI NON SEXEM

DATE DES DERNIERES REGLES

--	--	--	--	--	--	--

INTERROGATOIRE

(Code Clinique)

Le code de l'appréciation de l'intensité du symptôme est le suivant :

0 = symptôme n'existe pas

2 = signe certainement présent

1 = signe douteux

3 = signe présent mais très intense

Dans le cas où le signe est présent mais très ancien, sans rapport avec la maladie actuelle ou en rapport avec une autre cause évidente mettre 1, 2 ou 3 après avoir mis NON.

Par exemple : - malade émotif mais l'ayant toujours été

- malade présentant une diarrhée mais atteint d'une vieille colite

- Le malade est-il IRRITABLE ?

Intensité du signe

L'information existe-t-elle ?

OUI NON

IRITE

0 1 2 3

IRITI

OUI NON

IRITA

- Le malade est-il EMOTIF ?

Intensité du signe

L'information existe-t-elle ?

OUI NON

EMOTV

0 1 2 3

EMOTI

OUI NON

EMOTA

- Le malade a-t-il des PALPITATIONS ?

Intensité du signe

L'information existe-t-elle ?

OUI NON

PALPE

0 1 2 3

PALPI

OUI NON

PALPA

- Le malade a-t-il des INSOMNIES ?

Intensité du signe

L'information existe-t-elle ?

OUI NON

INSOM

0 1 2 3

INSOI

OUI NON

INSOA

- Le malade a-t-il des TREMBLEMENTS PERMANENTS ?

Intensité du signe

L'information existe-t-elle ?

OUI NON

TREMB

0 1 2 3

TREMI

OUI NON

TREMA

- Le malade a-t-il MAIGRI ?

Durée de la période d'amaigrissement en mois

Valeur de l'amaigrissement en kilog.

L'information existe-t-elle ?

OUI NON

AMAIG

AMAID

AMAIR

OUI NON

AMAIA

- L'amaigrissement a-t-il été volontaire ?

L'information existe-t-elle ?

OUI NON

AMAGV

OUI NON

AMAGA

- Le malade offre-t-il une modification de son APPETIT ?

Intensité du signe : 0 inchangé - 1 diminué - 2 augmenté - 3 boulimie

L'information existe-t-elle ?

OUI NON

APETI

0 1 2 3

APETH

OUI NON

APETA

- Le malade offre-t-il une AUGMENTATION DE LA SOIF ?

Intensité du signe

L'information existe-t-elle ?

OUI NON

SOMOD

0 1 2 3

SOMOI

OUI NON

SOMOA

- Le malade TRANSPIRE-T-IL exagérément ?
Intensité du signe
L'information existe-t-elle ?
- Le malade offre-t-il de la THERMOPHOBIE ?
Intensité du signe
L'information existe-t-elle ?
- Le malade est-il DIARRHEIQUE ?
Intensité du signe
L'information existe-t-elle ?
- Le malade offre-t-il une tendance à L'HYPERTHERMIE ?
PREDOMINANCE le matin, midi ou soir ? (code ci-contre)
Intensité du signe
L'information existe-t-elle ?
- Le malade offre-t-il de L'ASTHENIE LE MATIN ?
Intensité du signe
L'information existe-t-elle ?
- Le malade offre-t-il de L'ASTHENIE après L'EFFORT ?
Intensité du signe
L'information existe-t-elle ?
- Le malade offre-t-il des SIGNES PSYCHIQUES nets ?
Intensité du signe
L'information existe-t-elle ?
- Préciser lesquels ?.....
(Préciser si le symptôme est ancien ou récent ?).....
.....
.....
- Le malade est-il FRILEUX ?
Intensité du signe
L'information existe-t-elle ?
- Le malade a-t-il AUGMENTÉ récemment DE POIDS ?
Intensité du signe
L'information existe-t-elle ?
- Le malade est-il CONSTIPE ?
Intensité du signe
L'information existe-t-elle ?
- Le malade a-t-il des PERTES DE MEMOIRE ?
Intensité du signe
L'information existe-t-elle ?

OUI	NON	SUEUR	<input type="text"/>
0	1	2	3
OUI	NON	SUEUI	<input type="text"/>
OUI	NON	SUEUA	<input type="text"/>
OUI	NON	TEFOD	<input type="text"/>
0	1	2	3
OUI	NON	TEFOI	<input type="text"/>
OUI	NON	TEFOA	<input type="text"/>
OUI	NON	DIARR	<input type="text"/>
0	1	2	3
OUI	NON	DIARI	<input type="text"/>
OUI	NON	DIARA	<input type="text"/>
OUI	NON	HYTHE	<input type="text"/>
ma	mi	so	
1	2	3	
OUI	NON	HYTHM	<input type="text"/>
0	1	2	3
OUI	NON	HYTHI	<input type="text"/>
OUI	NON	HYTHA	<input type="text"/>
OUI	NON	ASTEM	<input type="text"/>
0	1	2	3
OUI	NON	ASTEI	<input type="text"/>
OUI	NON	ASTEA	<input type="text"/>
OUI	NON	ASTIE	<input type="text"/>
0	1	2	3
OUI	NON	ASTII	<input type="text"/>
OUI	NON	ASTIA	<input type="text"/>
OUI	NON	PSYCH	<input type="text"/>
0	1	2	3
OUI	NON	PSYCI	<input type="text"/>
OUI	NON	PSYCA	<input type="text"/>

OUI	NON	FRILO	<input type="text"/>
0	1	2	3
OUI	NON	FRILI	<input type="text"/>
OUI	NON	FRILA	<input type="text"/>
OUI	NON	OGPOD	<input type="text"/>
0	1	2	3
OUI	NON	OGPOI	<input type="text"/>
OUI	NON	OGPOA	<input type="text"/>
OUI	NON	CONST	<input type="text"/>
0	1	2	3
OUI	NON	CONSI	<input type="text"/>
OUI	NON	CONSA	<input type="text"/>
OUI	NON	PERME	<input type="text"/>
0	1	2	3
OUI	NON	PERMI	<input type="text"/>
OUI	NON	PERMA	<input type="text"/>

- Le malade a-t-il des CRAMPES DES MEMBRES INFERIEURS ?	OUI	NON	CRAME	<input type="text"/>		
Intensité du signe	0	1	2	3	CRAMI	<input type="text"/>
L'information existe-t-elle ?	OUI	NON	CRAMA	<input type="text"/>		
- IMPUISSANCE sexuelle MASCULINE ?	OUI	NON	PUSEX	<input type="text"/>		
Intensité du signe	0	1	2	3	PUSEI	<input type="text"/>
L'information existe-t-elle ?	OUI	NON	PUSEA	<input type="text"/>		
- LA LIBIDO MASCULINE a-t-elle disparue ?	OUI	NON	LIBIH	<input type="text"/>		
Intensité du signe	0	1	2	3	LIBII	<input type="text"/>
L'information existe-t-elle ?	OUI	NON	LIBIA	<input type="text"/>		
- Existe-t-il une FRIGIDITE FEMININE ayant débuté à peu près en même temps que l'affection thyroïdienne ?	OUI	NON	FRIGF	<input type="text"/>		
Intensité du signe	0	1	2	3	FRIGI	<input type="text"/>
L'information existe-t-elle ?	OUI	NON	FRIGA	<input type="text"/>		
- LA LIBIDO FEMININE a-t-elle disparue ?	OUI	NON	LIBFE	<input type="text"/>		
Intensité du signe	0	1	2	3	LIBFI	<input type="text"/>
L'information existe-t-elle ?	OUI	NON	LIBFA	<input type="text"/>		
- La malade a-t-elle DEJA ETE REGLEE ?	OUI	NON	REGLI	<input type="text"/>		
L'information existe-t-elle ?	OUI	NON	REGLA	<input type="text"/>		
- La malade a-t-elle des périodes récentes D'AMENORRHEE ?	OUI	NON	AMENO	<input type="text"/>		
L'information existe-t-elle ?	OUI	NON	AMENA	<input type="text"/>		
- La malade a-t-elle eu des périodes récentes de METRORRAGIES ?	OUI	NON	METRO	<input type="text"/>		
L'information existe-t-elle ?	OUI	NON	METRA	<input type="text"/>		
- CYCLE MENSTRUEL						
Ex : 30.5 = cycle de 30 jours						
durée des règles de 5 jours						
- La malade est-elle COMPLETEMENT MENOPAUSEE ?	OUI	NON	MENOF	<input type="text"/>		
L'information existe-t-elle ?	OUI	NON	MENOA	<input type="text"/>		
Age de la ménopause :						
- La malade a-t-elle subi une CASTRATION ?	OUI	NON	CASTR	<input type="text"/>		
L'information existe-t-elle ?	OUI	NON	CASTA	<input type="text"/>		
- La malade est-elle EN COURS DE MENOPAUSE ?	OUI	NON	MENFE	<input type="text"/>		
L'information existe-t-elle ?	OUI	NON	MENFA	<input type="text"/>		
- La malade a-t-elle eu des ACCOUCHEMENTS ?	OUI	NON	ACCOU	<input type="text"/>		
L'information existe-t-elle ?	OUI	NON	ACCOA	<input type="text"/>		
- Préciser le nombre :						
- Nombre d'ENFANTS VIVANTS ?			EMVIV	<input type="text"/>		
L'information existe-t-elle ?	OUI	NON	EMVIA	<input type="text"/>		
- Nombre de FAUSSES COUCHES ?			FOSCO	<input type="text"/>		
L'information existe-t-elle ?	OUI	NON	FOSCA	<input type="text"/>		
- Date de la dernière grossesse						

E X A M E N

- Poids (en kg arrondis) :

POIDS

- Taille (en cm arrondis) :

TAILL

LE FACIES

- Le COMPORTEMENT général est-il ANORMAL ?

OUI NON COMGE

Intensité du signe - trop calme 1 - normal 0
- trop agité 2

0 1 2 COMGI

L'information existe-t-elle ?

OUI NON COMGA

- Le malade a-t-il un FACIES HYPOTHYROIDIEN ?

OUI NON FAIPO

Intensité du signe

0 1 2 3 FAIPI

L'information existe-t-elle ?

OUI NON FAIPA

- Le malade présente-t-il un OEDEME DU VISAGE ?

OUI NON OEVIS

Intensité du signe

0 1 2 3 OEVI

L'information existe-t-elle ?

OUI NON OEVI

- Le malade a-t-il une LANGUE EPAISSE ?

OUI NON MAGLO

L'information existe-t-elle ?

OUI NON MAGLA

- Le malade offre-t-il une EXOPHTALMIE ?

OUI NON EXOPH

L'information existe-t-elle ?

OUI NON EXOPA

Mesure à l'exophthalmomètre en mm O.D. O.G.

- Existe-t-il un REGARD BRILLANT ?

OUI NON REBRI

L'information existe-t-elle ?

OUI NON REBRA

- Existe-t-il une RETRACTION de la PAUPIERE SUPERIEURE ?

OUI NON REPAU

L'information existe-t-elle ?

OUI NON REPAN

- Existe-t-il un OEDEME DES PAUPIERES ?

OUI NON OEPAU

L'information existe-t-elle ?

OUI NON OEPA

- Existe-t-il une INFLAMMATION DES CONJONCTIVES ?

OUI NON INCON

L'information existe-t-elle ?

OUI NON INCOA

- Existe-t-il de la PHOTOPHOBIE ?

OUI NON PHOTO

L'information existe-t-elle ?

OUI NON PHOTA

- Existe-t-il une PARALYSIE OCULAIRE ?

OUI NON PAROC

L'information existe-t-elle ?

OUI NON PAROA

- Préciser lesquelles :

.....

EXAMEN ISOTOPIQUE

TAUX DE FIXATION

- Tx de FIXATION thyroïdienne à la 6^e HEURE augmenté ? (sup. à 37%)
Tx de fixation thyroïdienne à la 6^e heure diminué ? (inf. à 19%)
Valeur du tx de fixation à la 6^e heure :
L'information existe-t-elle ?
- Tx de FIXATION thyroïdienne à la 24^e HEURE augmenté ? (sup. à 54%)
Tx de fixation thyroïdienne à la 24^e heure diminué ? (inf. à 30%)
Valeur du tx de fixation à la 24^e heure :
L'information existe-t-elle ?
- Tx de FIXATION à la 48^e HEURE augmenté (sup. à 56%)
Tx de fixation à la 48^e heure diminué (inf. à 34%)
Valeur du tx de fixation à la 48^e heure :
L'information existe-t-elle ?

OUI	NON	TF6HP	<input type="text"/>
OUI	NON	TF6HM	<input type="text"/>
<input type="text"/>		TF6HV	<input type="text"/>
OUI	NON	TF6HA	<input type="text"/>
OUI	NON	TF24P	<input type="text"/>
OUI	NON	TF24M	<input type="text"/>
<input type="text"/>		TF24V	<input type="text"/>
OUI	NON	TF24A	<input type="text"/>
OUI	NON	TF48P	<input type="text"/>
OUI	NON	TF48M	<input type="text"/>
<input type="text"/>		TF48V	<input type="text"/>
OUI	NON	TF48A	<input type="text"/>

BLANQUET

- F.I DU BLANQUET 24 h augmenté ? (sup. à 92 %)
F.I du Blanquet 24 h diminué ? (inf. à 64 %)
Valeur du F.I Blanquet 24 h :
L'information existe-t-elle ?
- F.II DU BLANQUET 24 h augmenté ? (sup. à 10%)
F.II du Blanquet 24 h diminué ? (inf. à 1%)
Valeur du F.II Blanquet 24 h :
L'information existe-t-elle ?
- F.III DU BLANQUET 24 h augmenté ? (sup. à 25%)
F.III du Blanquet 24 h diminué ? (inf. à 5%)
Valeur du F.III Blanquet 24 h :
L'information existe-t-elle ?
- Un BLANQUET-MEYNIEL a-t-il été pratiqué à la 48^eme HEURE ?
L'information existe-t-elle ?
- Valeur du rapport PBI 131/plasma/24 h
- Valeur du rapport PBI 131/plasma/48 h

OUI	NON	B124P	<input type="text"/>
OUI	NON	B124M	<input type="text"/>
<input type="text"/>		B124V	<input type="text"/>
OUI	NON	B124A	<input type="text"/>
OUI	NON	B224P	<input type="text"/>
OUI	NON	B224M	<input type="text"/>
<input type="text"/>		B224V	<input type="text"/>
OUI	NON	B224A	<input type="text"/>
OUI	NON	B324P	<input type="text"/>
OUI	NON	B324M	<input type="text"/>
<input type="text"/>		B324V	<input type="text"/>
OUI	NON	B324A	<input type="text"/>
OUI	NON	BL48M	<input type="text"/>
OUI	NON	BL48A	<input type="text"/>
<input type="text"/>		PP24V	<input type="text"/>
<input type="text"/>		PP48V	<input type="text"/>

SERVICE CENTRAL DES RADIOISOTOPES
DU CENTRE HOSPITALO-UNIVERSITAIRE DE MARSEILLE
HOPITAL EST - GROUPEMENT HOSPITALIER DE LA TIMONE (PR. H. ROUX)

EXPLORATION THYROIDIENNE

RESUME DU DOSSIER
=====

EDITE PAR L'ORDINATEUR DU SERVICE DE BIOMATHÉMATIQUES, STATISTIQUES ET INFORMATIQUE MÉDICALE DE LA FACULTÉ DE MÉDECINE.

** IDENTIFICATION **
=====

5994 NOM: ~~CHIRURGIE~~ ~~CHIRURGIE~~ NOM D'ÉPOUSE: ~~MARIE-LOUISE~~

AGE: 63 ANS; SEXE: FEMININ

DATE DE CONSULTATION: 19 / 10 / 1978

***** MÉDICAMENT PERTURBATEUR *****

*** TRAITEMENT À VISÉE THYROIDIENNE ***
CHIRURGIE

** PRINCIPAUX SYMPTÔMES PRÉSENTES PAR LE MALADE **
=====

* AMAIGRISSEMENT	NON	* PRISE DE POIDS	OUI
APPETIT AUGMENTÉ	NON	APPELIT DIMINUÉ	NON
PALPITATIONS	NON		
TENSION NERVEUSE	NON	CRAMPES DES MEMBRES INFÉRIEURS	OUI
INSOMNIE	NON		
* THERMOPHOBIE	NON	FRILLOSITÉ	OUI
* HYPERSUADATION DE LA SOIF	NON	CHUTE DES CHEVEUX	NON
DIARRHÉE	NON	ONGLES CASSANTS	NON
POULS SUPÉRIEUR À 90	NON	CONSTIPATION	OUI
ERETHISME CARDIAQUE	NON	POULS INFÉRIEUR À 80	OUI
MAINS CHAUDES	NON		
MAINS MOITES	NON	MAINS FROIDES	OUI
* THORAX MOITE	NON	MAINS SÈCHES	OUI
AGITATION OBJECTIVE	NON	* PEAU SÈCHE, SQUAMEUSE	OUI
TREMBLEMENT	NON	APATHIE OBJECTIVE	NON
SIGNE DU TABOURET	NON	* VOIX RAUQUE ET NASONNÉE	OUI
ARYTHMIE COMPLÈTE	NON	RAREFACTION DE LA G. DU SOURCIL	NON
VITILIGO	NON	OEDEME DU VISAGE	NON
MECHE BLANCHE	NON		
MYXOEDEME PRETIBIAL	NON		
EXOPHTALMIE	NON		
RETRACTION FAUPIÈRE	NON		

OUI

OEDEME DES PAUPIERES

EXAMEN DU CORPS THYROIDE :

PAS DE GOITRE

EXAMENS PARACLINIQUES

EXAMENS	RESULTATS	VALEURS NORMALES
- IODEMIE PROTIIDIQUE	4.20 MCG/100ML	4 A 8
- IODE HORMONAL DU PLASMA	3.80 MCG/100ML	3.5 A 7.5
- CHOLESTEROLEMIE	3.60 G/LITRE	1.5 A 2.5
- LIPIDEMIE	9.00	
- GLOBULES ROUGES	4.40 MILLIONS	4 A 5.5
- GLOBULES BLANCS	6300.00 UNITES	
- POLYNUCLEAIRES	65.00 %	
- LYMPHOCYTES	33.00 %	
- V.S. A LA 1ERE HEURE	18.00	
- TAUX DE FIXATION A LA 6EME H.	2.00 %	15 A 25
- TAUX DE FIXATION A LA 24EME H.	5.00 %	25 A 54
- T4 TEST (TETRASORE)	1.40 MCG/100ML	5.5 A 11.5
- T3 TEST (TRIOSORE)	23.00 %	25 A 35
- I.T.L	.30	1.5 A 3.4

** SCINTIGRAPHIE **

SCINTIGRAPHIE PRATIQUEE AVEC TC99M
FIXATION THYROIDIENNE EN POSITION NORMALE
SURFACE CIRCONSCRITE PAR LE CONTOUR= 11 CM2
LA SURFACE DE FIXATION EST PLUS GRANDE A DROITE QU'A GAUCHE
FIXATION GLOBALEMENT AUSSI DENSE A DROITE QU'A GAUCHE

N.B.: POUR LES MALADES HOSPITALISES, L'EXAMEN DOIT ETRE TARIFIE Z20 + 20 = Z40 AU CODE 1770.

5994 NOUVELLES FONCTIONS LOGIQUES

=====

FONCTION NO. 3 VERIFIEE. (MIN NO. 92, EFF = 10)

DIAGNOSTIC D'AFFINEMENT

=====

SUSPICION D'HYPOTHYROIDIE FRUSTE
PRATIQUER UN DOSAGE DE LA TSH POUR CONFIRMER LE DIAGNOSTIC.
HYPOTHYROIDIE SECONDAIRE AU TRAITEMENT D'UNE HYPER.
EXCES PONDERAL APPRECIABLE

DIAGNOSTIC SERVICE

=====

DIAGNOSTIC D'ENVOI :

DIAGNOSTIC INITIAL :

DIAGNOSTIC FINAL :

PERTURBATION : CHIRURGIE
FONCTIONNEL : HYPOTHYROIDIE FRANCHE
MORPHOLOGIQUE : HYPOPLASIE
SCINTIGRAPHIQUE: HOMOGENE
MECANISME : IATROGENE

CHEMIN PARCOURU DANS L'ORGANIGRAMME :

1019	1034	1058	1073	1074	1089	1090
1105	1118	1132	1144	1159	1175	1148
1149	1150					


```

NOM DU SYSTEME D AXIOMES A TRAITER .....THYROIDE
NBE DE PAS ..10
TRAVAIL SUR 1 SEUL SYSTEME(1) OU SUR TOUS LES SYSTEMES DU GROUPE(2) :
NO DU GROUPE _ 0:DEFILEMENT _ 1.40:POSITIONNEMENT ..0
1 HYPO
2 NORMAL
3 HYPER
4 THYREO
NUM.CHOISI?= ..2
SYST.D AXIOMES CHOISIS NORMAL
NEGATION A DEMONTRER
.-S11/-S7/-S12/-S6*
LES SIGNES PRESENTES PAR LE PATIENT ENTRAINENT LE DIAGNOSTIC DE NORMAL
NOUVEAU TRT(=0) OU FIN? ..0
NOM DU SYSTEME D AXIOMES A TRAITER .....THYROIDE
NBE DE PAS ..10
TRAVAIL SUR 1 SEUL SYSTEME(1) OU SUR TOUS LES SYSTEMES DU GROUPE(2) :
NO DU GROUPE _ 0:DEFILEMENT _ 1.40:POSITIONNEMENT ..2
SYST.D AXIOMES CHOISIS NORMAL
NEGATION A DEMONTRER
.-S11/-S7/-S12*
LES SIGNES PRESENTES PAR LE PATIENT N ENTRAINENT PAS LE DIAGNOSTIC DE
NOUVEAU TRT(=0) OU FIN? ..0
NOM DU SYSTEME D AXIOMES A TRAITER .....S11MMMMM
LE SYST.CHOISI N EXISTE PAS
NOM DU SYSTEME D AXIOMES A TRAITER .....THYROIDE
NBE DE PAS ..10
TRAVAIL SUR 1 SEUL SYSTEME(1) OU SUR TOUS LES SYSTEMES DU GROUPE(2) :
NO DU GROUPE _ 0:DEFILEMENT _ 1.40:POSITIONNEMENT ..2
SYST.D AXIOMES CHOISIS NORMAL
NEGATION A DEMONTRER
.-S11/-S7/-S12/-S6*
LES SIGNES PRESENTES PAR LE PATIENT N ENTRAINENT PAS LE DIAGNOSTIC DE
NOUVEAU TRT(=0) OU FIN? ..0
NOM DU SYSTEME D AXIOMES A TRAITER .....THYROIDE
NBE DE PAS ..10
TRAVAIL SUR 1 SEUL SYSTEME(1) OU SUR TOUS LES SYSTEMES DU GROUPE(2) :
NO DU GROUPE _ 0:DEFILEMENT _ 1.40:POSITIONNEMENT ..0
1 HYPO
2 NORMAL
3 HYPER
4 THYREO
NUM.CHOISI?= ..3
SYST.D AXIOMES CHOISIS HYPER
NEGATION A DEMONTRER
.-S13/S16/S17/S11/-S2/S4/-S6*
LES SIGNES PRESENTES PAR LE PATIENT ENTRAINENT LE DIAGNOSTIC DE HYPER
NOUVEAU TRT(=0) OU FIN? ..0
NOM DU SYSTEME D AXIOMES A TRAITER .....THYROIDE
NBE DE PAS ..10
TRAVAIL SUR 1 SEUL SYSTEME(1) OU SUR TOUS LES SYSTEMES DU GROUPE(2) :
NO DU GROUPE _ 0:DEFILEMENT _ 1.40:POSITIONNEMENT ..0
1 HYPO
2 NORMAL
3 HYPER
4 THYREO
NUM.CHOISI?= ..4
SYST.D AXIOMES CHOISIS THYREO
NEGATION A DEMONTRER
.-S12/S14/S18/-S11.S13/-S1*
LES SIGNES PRESENTES PAR LE PATIENT ENTRAINENT LE DIAGNOSTIC DE THYREO
NOUVEAU TRT(=0) OU FIN? ..0

```


NOM DU SYSTEME D AXIOMES A TRAITERTHYROIDE
 NBE DE PAS ..10
 TRAVAIL SUR 1 SEUL SYSTEME(1) OU SUR TOUS LES SYSTEMES DU GROUPE(2) : .2
 SYSTEME TRAITE: HYPO
 NEGATION A DEMONTRER
 .S1/S3/S5/S7/S9*
 LES SIGNES PRESENTES PAR LE PATIENT ENTRAINENT LE DIAGNOSTIC DE HYPO
 SYSTEME TRAITE: NORMAL
 NEGATION A DEMONTRER
 .S1/S3/S5/S7/S9*
 LES SIGNES PRESENTES PAR LE PATIENT N ENTRAINENT PAS LE DIAGNOSTIC DE NOI
 SYSTEME TRAITE: HYPER
 NEGATION A DEMONTRER
 .S1/S3/S5/S7/S9*
 LES SIGNES PRESENTES PAR LE PATIENT N ENTRAINENT PAS LE DIAGNOSTIC DE HYF
 SYSTEME TRAITE: THYREO
 NEGATION A DEMONTRER
 .S1/S3/S5/S7/S9*
 LES SIGNES PRESENTES PAR LE PATIENT N ENTRAINENT PAS LE DIAGNOSTIC DE THY
 NOUVEAU TRT(=0) OU FIN? ..0
 THYREO

NOM DU SYSTEME D AXIOMES A TRAITERTHYROIDE
 NBE DE PAS ..10
 TRAVAIL SUR 1 SEUL SYSTEME(1) OU SUR TOUS LES SYSTEMES DU GROUPE(2) : ?
 SYSTEME TRAITE: HYPO
 NEGATION A DEMONTRER
 .-S11/-S6/-S12/-S7*
 LES SIGNES PRESENTES PAR LE PATIENT N ENTRAINENT PAS LE DIAGNOSTIC DE HYI
 SYSTEME TRAITE: NORMAL
 NEGATION A DEMONTRER
 .-S11/-S6/-S12/-S7*
 LES SIGNES PRESENTES PAR LE PATIENT ENTRAINENT LE DIAGNOSTIC DE NORMAL
 SYSTEME TRAITE: HYPER
 NEGATION A DEMONTRER
 .-S11/-S6/-S12/-S7*
 LES SIGNES PRESENTES PAR LE PATIENT N ENTRAINENT PAS LE DIAGNOSTIC DE HYI
 SYSTEME TRAITE: THYREO
 NEGATION A DEMONTRER
 .-S11/-S6/-S12/-S7*
 LES SIGNES PRESENTES PAR LE PATIENT N ENTRAINENT PAS LE DIAGNOSTIC DE THY
 NOUVEAU TRT(=0) OU FIN? ..FI
 THYREO

NOM DU SYSTEME D AXIOMES A TRAITERTHYROIDE
 NBE DE PAS ..10
 TRAVAIL SUR 1 SEUL SYSTEME(1) OU SUR TOUS LES SYSTEMES DU GROUPE(2) : .2
 SYSTEME TRAITE: HYPO
 NEGATION A DEMONTRER
 .S13/S15/S17/S11*
 LES SIGNES PRESENTES PAR LE PATIENT N ENTRAINENT PAS LE DIAGNOSTIC DE HYF
 SYSTEME TRAITE: NORMAL
 NEGATION A DEMONTRER
 .S13/S15/S17/S11*
 LES SIGNES PRESENTES PAR LE PATIENT N ENTRAINENT PAS LE DIAGNOSTIC DE NOI
 SYSTEME TRAITE: HYPER
 NEGATION A DEMONTRER
 .S13/S15/S17/S11*
 LES SIGNES PRESENTES PAR LE PATIENT ENTRAINENT LE DIAGNOSTIC DE HYPER
 SYSTEME TRAITE: THYREO
 NEGATION A DEMONTRER
 .S13/S15/S17/S11*
 LES SIGNES PRESENTES PAR LE PATIENT N ENTRAINENT PAS LE DIAGNOSTIC DE THY
 NOUVEAU TRT(=0) OU FIN? ..FI
 THYREO

TRAVAIL DEMANDESTOP

BIBLIOGRAPHIE

- [1] ADIN M. (Groupe de travail), "L'informatique médicale", Collection "Que sais-je", éd. Presses Universitaires de France.
- [2] ANDERSON, "An Introduction to Multivariate Statistical Analysis", Wiley Publications in Statistics Ed.
- [3] F. BEGON, R. TREMOLIERES, "Elaboration d'un modèle logique du diagnostic", Supplément au bulletin de l'I.R.I.A., n°9, octobre 1971.
- [4] F. BEGON, R. TREMOLIERES, "Présentation et illustration d'un modèle logique du diagnostic", Notes de l'Institut d'Informatique et d'Automatique de Rocquencourt.
- [5] CHIEN-LIANG CHANG, Richard CHAR-TUNG LEE, "Symbolic Logic and Mechanical Theorem Proving", Academic Press.
- [6] KAUFMANN A., "Introduction à la Théorie des sous-ensembles flous - Tome I : Eléments théoriques de base", Ed. Masson.
- [7] LEBART L., FENELON J.P., "Statistique et Informatique appliquées", Ed. DUNOD, pp. 240-251.
- [8] LEFEVRE J., "Introduction aux analyses statistiques multidimensionnelle".
- [9] NORA S., MINC A., "L'informatisation de la Société", Ed. Points "La documentation française".
Document contributif n° 4 : FUNCK J.L., BRENTANO, "L'introduction de l'informatique dans les activités médicales et de santé".
- [10] ROUX M., BARRE A., "Contribution à l'étude de l'aide au diagnostic", Revue Informatique médicale, 1971.
- [11] ROUX M., GOUVERNET J., AURRAN Y., "Application au dossier thyroïdien de méthodes de traitement des variables numériques continues et de méthodes de traitement des variables logiques", Revue Informatique médicale, 3(2), pp. 131-136, 1972.
- [12] ROUX M., BARRE A., "Méthode étape par étape", Revue Informatique médicale, 3(2), pp. 137-141, 1972.
- [13] ROUX M., FIESCHI M., SAMBUC, GOUVERNET J., SAN MARCO J.L., SIMONIN R., "Aide à la décision en médecine", Pathologie - Biologie, 26, n°2, pp. 107-115.
- [14] ROUX M., "Sur un modèle de gestion et d'exploitation du dossier médical thyroïdien par les calculateurs digitaux", Thèse de doctorat en Médecine, Faculté de Médecine, Marseille.

- [15] ROUX M., "Contribution à l'étude de l'aide au diagnostic des affections thyroïdiennes", *Revue Informatique médicale*, 4(3), pp. 265-273, 1972.
- [16] ROUX M., "Le Dossier Médical : aspects cliniques, aspects techniques, aspects mathématiques", publié par la Section Biomathématiques Statistiques et Informatique médicale de la Faculté de Médecine de Marseille.
- [17] ROUX M., GOUVERNET J., FIESCHI M., JOUBERT M., LOMBARD, CORSON A., "Mise en place d'un système informatique de dossiers médicaux : le système SYCVAR", *IRIA 'Journées Informatiques Médicales de Toulouse'*, 1975.
- [18] ROUX M., GOUVERNET J., "Sur un mode de gestion du dossier médical : les dossiers d'exploitation fonctionnelle, thyroïdienne et cardio-respiratoire", *Revue Informatique médicale*, (3), septembre 1970.
- [19] ROUX M., MORELLI P., "Sur un système de mise à jour d'un dossier médical", *'Journées Informatiques médicales'*, Toulouse, 1970.
- [20] SAMBUC R., "Fonctions Φ -floues, application à l'aide au diagnostic en pathologie thyroïdienne", Thèse de doctorat d'état en Médecine, Faculté de Médecine, Marseille.
- [21] SANCHEZ E., SAMBUC R., "Relations floues - Fonctions Φ -floues", *Medical Data Processing*, Taylor and Francis LTD (London).
- [22] SANCHEZ E., "Resolution of composite fuzzy relation equations", *Information and Control*, Décembre 1975.
- [23] SANCHEZ E., AURRAN Y., BARRE A., GOUVERNET J., "Inférences par compositions en logique floue. Application à l'aide à la décision en médecine", *'Journées Informatiques médicales'*, Toulouse, 1978.

BUMP



0 0 3 2 1 2 9 3 3

*FM B16/1979/16

